H I L G A R D I A

A Journal of Agricultural Science Published by the California Agricultural Experiment Station

VOL. 28 JUNE, 1959 No. 18

MODERN SENSORY METHODS OF EVALUATING WINE[1]

M. A. AMERINE, E. B. ROESSLER, and F. FILIPELLO[2]

FOREWORD

WITH THE GROWING CONSUMPTION of wine in this country and the rapidly developing interest of the general public in wine appreciation, maintenance of uniform quality and improvement in quality have become matters of increasing importance to California wine makers.

At present our only method of distinguishing quality is by sensory examination—visual observation, smelling, and tasting—and this is the province of professionals, calling for skill, training, and experience. To some extent objective tests can be applied to supplement and support sensory judgments, but quality ratings still rest largely on the estimates of expert tasters.

While much has been written about the various aspects of wine judging, the information is scattered and not always easily available. Furthermore, because of their subjective nature, standards tend to vary, and a major problem of wine tasting has been to achieve, by various types of testing, uniform and reliable judgments.

The purpose of this publication is to provide a complete guide to the sensory examination of all kinds of wines. The standards proposed herein are generally accepted measures of quality, and an attempt has been made, insofar as possible, to provide objective criteria that will be useful both to amateur and experienced tasters.

The first section is devoted to a discussion of the senses and the way they function, since this is considered fundamental to their intelligent utilization. The chemical characteristics of wines are discussed under color, smell, and taste, and the available data on their detection or differentiation are summarized. A list of words for purely descriptive work is given in connection with each of these. A brief description of the main types of wine produced in California then follows. Various types of tasting for different purposes are suggested. Difference testing is discussed, and the statistical procedures for determining significance are outlined. Finally, scoring and ranking pro-

[1] Received for publication April 30, 1957.

[2] Mr. Amerine is Professor of Enology and Enologist in the Experiment Station, Davis; Mr. Roessler is Professor of Mathematics and Statistician in the Experiment Station, Davis; Mr. Filipello is Associate Specialist in Food Technology, Experiment Station, Davis.

cedures are considered, and the more complicated statistical methods required are presented. A method for selecting panels by a sequential statistical procedure is also given. It should be emphasized that the statistical procedures are given only in outline or summary form. For a more complete analysis of these, the reader is referred to statistical treatises such as Alder and Roessler (1951, 1954), Bradley *et al.* (1950), Fisher and Yates (1953), Snedecor (1946), or Yule and Kendall (1950).

I. INTRODUCTION

No other food product has a longer history of quality evaluation than wine. Homer's Pramnian and Maronean, Pliny's Falernian, and Horace's Chios and Lesbos are examples of wines that were famous long before, or at the beginning of, the Christian era. Their fame obviously rested on some sort of subjective organoleptic examination, or at least on comparisons between wines.

Such early classifications based on quality evaluation led directly to the modern systems of varietal or regional classification. In some cases the regions have well-defined subregions and even elaborately subdivided vineyards, each subdivision producing wine of a slightly different style and quality. Variability in composition of grapes or in processing is primarily responsible for the differences in the composition, and hence in type and quality, of wines. The many varieties and blends of grapes used, the wide range of climatic zones in which the vines are grown, and the special climatic conditions of the particular vintage season, as well as the cultural practices of the individual producer, all help to explain the great diversification. Various production and aging procedures also contribute to the type and quality distinctions that have arisen.

The existence of specialized wine names differentiating vineyards and years is good evidence that quality is a recognizable and distinguishable entity in wines. The perishability of wine was probably one factor that first led men to taste it critically. The step between wine and vinegar was never very great in the primitive wine industry, and, lacking chemical procedures, men relied on sensory tests to detect the difference. However, the aesthetic experience was also an early aspect of wine appraisal. Horace wrote of wine as if its appreciation were an aesthetic matter, which indeed it is to the experienced wine consumer. The quality of the wine was apparently a matter comparatively separate from its popular appeal. An interesting, though perhaps overgenerous appraisal of Roman wines has been made by Allen (1931).

Wine tasting has, until recently, been performed by professionals. Considerable quantities of wines are still purchased by skilled wine brokers who base their selections upon their taste evaluations. Within a limited range of wine types, such tasters are probably authoritative and can undoubtedly duplicate their results. Amerine and Feduchy (1954) found this to be true with relatively untrained Spanish tasters. Papakyriakopoulos and Amerine (1956) also demonstrated this for moderately trained American tasters. They suggested that quality must be a relatively fixed concept, although varying between tasters.

Tasters may differ markedly from each other in quality appraisal, and they may do so quite consistently according to Filipello (1957). This appears to be due to differences in experience among the tasters and to individual aesthetic preferences. However, the quality judgment of the untrained taster is even more frequently unreliable because it may change with experience. Even the aesthetic preferences of the trained critic *may* be modified or expanded by contact with new ideas or by the refinement of continued tasting of a certain type of wine.

There are two types of quality evaluation, each involving different concepts and methods and presenting its special problems. One—and the one of primary concern in this publication is the discrimination between two wines. This discrimination may involve assigning quality ratings to two or more wines, or it may be used simply to detect differences between wines.

In contrast to such critical discrimination is consumer preference. The first objective of consumer research is to determine which of two or more wines is preferred by a selected group of consumers; its next concern is to predict what the preference of the consumer may be after repeated trials or further experience. It is obvious that taste preferences involving a changing standard introduce difficult problems of testing and analysis. From indications of present preferences, we cannot draw infallible conclusions about the future preferences of a group of consumers. To our knowledge, this has not yet been successfully analyzed. Of a particular commodity, the quantity consumed is only one measure of quality and an inadequate one. Because of the difficulty of defining the quantity measured, an analysis of this type of tasting is not included. We shall, however, make some comments on experiments which seem to show how specific components of wine influence consumer acceptance. For a good analysis of the problems of consumer acceptance and new product development see Anon. (1958).

Although most producers will grant that quality does exist, they argue that it is unimportant, since the "average" consumer cannot or does not recognize it. This argument is fallacious on more than one count. In the first place, the "average" consumer is merely a statistic. Furthermore, the trend with other products must be considered. To be a reliable statistic, the "average" consumer needs to be defined not only with respect to the group or class of the survey (that is, the product being tested) but also with respect to other groups.

Some of the public do recognize quality, although economic limitations may prevent them from purchasing freely. The inexperienced consumer, on the other hand, may be unduly influenced by trade promotion ideas based on the reputation of traditional types of wines.

Chamberlain (1955) has shown that introduction of a new product (or a new form of an old product) into the American economy involves two simultaneous movements: 1) a gradual reduction in price as mass production takes over; 2) an increase in the number of models or types as producers develop a complete selling line ranging from cheap to expensive items. Analyses of this concept with respect to wines might reveal some interesting possibilities.

On the other hand, Marteau (1953) has reported that with the increasing industrialization of the French wine industry has come a tendency to reduce the diversity of wine types produced. He also notes legislation that fixes

limits for certain constituents in order to prevent poor wines from reaching the market. How far such legislation can standardize practice is questionable. Legislation fixing the varieties permitted, production, and composition of typical French wines is another recent development. Marteau comments, correctly, that our knowledge of analytical chemistry is clearly insufficient to differentiate these "typical" wines. At present our only method of distinguishing quality is by tasting, although certain chemical and physical properties of wines can certainly be employed to supplement or support sensory judgments.

The tendency toward standardization and government control of *some* of the factors governing quality has been evident in the California wine industry since Repeal. Neither movement has reached its ultimate development. Our interest in these trends is as they relate to quality control because they indicate that a quality range in wines will be a permanent part of the California wine industry. As yet, however, a definitive relationship between the composition of a wine and its consumer acceptance or quality is certainly far from being achieved.

There are other indications that recognition of quality is a permanent part of the wine industry. The European wine countries where this is best demonstrated are France and Germany. In these countries, quality differentiation has been carried to the point that price is largely, but not entirely, established on the basis of quality. The wide difference in quality due to climatic variations from one year to the next makes this more noticeable. Reputation and past performance cannot raise the price of a famous wine in a bad year—not even for the less-experienced American market, although there seem to be some exceptions.

The importance of sensory examination of foods has increased rapidly in the 20th century. Prior to 1900, the consumer was often also a producer. The producers then made their own consumer-acceptance tests—on themselves. Even where this was impossible, the tradition of the consumer-as-producer remained deeply ingrained in the thinking of management. Later, however, the producer could no longer be considered a consumer. The plant grew too large. Many types of products were produced. Trained technicians insured the sanitation and uniformity of the product—often with scant attention to palatability. This situation applies not only to the food industries in general but to the wine industry in particular. For the reasons given above, the recognition of quality is now becoming an important aspect of these industries. In many food industries, however, uniformity is apparently still regarded as the most important aspect of quality.

However, even if quality *per se* were not an important requirement for a food industry, sensory examination would still be important. Uniformity *is* one of the necessary aspects of standard competitive foods. This means uniformity not only of chemical attributes but also of the gustatory and olfactory components of the food product. Slight differences in odor may result in changes in food preferences or acceptance. The whiskey industry has found it particularly valuable to use sensory panels as controls on the uniformity of old and new blends. It should also be noted that the triangular taste test system (see p. 526)—a powerful tool for *detecting differences*—was originally developed in Scandinavian breweries.

Organoleptic examination is thus a necessary part of the production process for all kinds of foods intended for human consumption—whether of low or high quality. At the level of the standard wine it must be employed to insure uniformity. With quality products it is required to determine the degree of quality, as a guide to both producers and consumers. We do not minimize the difficulty of defining "quality," but we do not find any inherent problems in its measurement. Admittedly, the techniques are often inadequate, but the philosophical and mathematical concepts seem adequate, even though they are difficult to apply and hence infrequently used.

Sensory examination of wines means using the senses of sight, smell, taste, and touch to determine differences between wines or to evaluate their quality. The term "sensory examination" is preferred to "tasting" because it is more general and does not imply that taste is the most important sense involved, as it certainly is not. For a recent and lucid discussion of the whole field of sensory examination of foods see Anon. (1958).

II. THE SENSES

Some knowledge of the senses involved is necessary for the intelligent use of sensory tests. However, the subject of the physiology of taste and smell is outside the scope of this publication. Only the basic essentials will be outlined. For more complete information, see Boring (1942), Geldhard (1953), Moncrieff (1951), Pfaffmann (1956), Skramlik (1926), or Tilgner (1957).

Sight

Since color and appearance are major aspects of wine quality, vision is the first sense used in the organoleptic examination. Color discrimination is very important in examining wines. In extreme cases a wine may be rejected on the basis of color and appearance alone, thus eliminating the other steps of evaluation. More commonly, color and appearance provide the first indication of a wine's superiority or inferiority. In addition, these two aspects of the evaluation process may often provide valuable clues that will aid the experienced taster's concentration when he comes to sniffing and tasting. The colors that we are concerned with in wines vary from some shade of red through orange to yellow and brown. Traces of green and violet are noted in certain wines.

Color is not a characteristic of objects but of the light from them which enters our eyes. Objects appear colored only because they reflect or transmit to our eyes radiant energy of certain predominant wave lengths in the visible spectrum. Different colors may be seen when various kinds of light illuminate an object. It has been shown, however, that chromatic perceptions are remarkably independent of the quality of the illumination.

The characteristics of light that define a color can be defined in terms of luminous transmittance (or reflection), dominant wave length, and purity. Instead of the awkward term "luminous transmittance," luminance, luminosity, or lightness are often used. These are the measurable, objective aspects of color. They correspond in a very general way to brightness, hue or tint, and saturation, which are attributes of the color sensation experienced by the viewer. There are other attributes of color perception, such as luster, glossi-

ness, transparency, and sparkle, but they are of lesser importance. Luminous reflection is the characteristic of the light reflected from pure white paper illuminated by a 100-watt lamp (all other factors being equal). Dominant wave length is that which appears to be the characteristic wave length of the color. Purity may be thought of as the degree to which the dominant wave length appears to predominate in the light. In comparing wines, care must be exercised when the brightness, dominant wave length, or saturation of the two samples is not the same.

Visual adaptation is very complex, for in the usual situation the retinal light pattern is continually changing. According to the Committee on Colorimetry of the Optical Society of America (1953), purity and luminance at first seem to decrease in response to a single unvarying chromatic stimulus incident upon the center of the retina. That is, the viewer's first reaction to the color is that it is less pure and less bright than it actually is. The eye quickly adapts, however, in one or two minutes reaching a constant perception or concept of the color's saturation. No important shift in hue occurs because of adaptation. The way that adaptation to one color affects the viewer's reaction to another is more complicated, but this is probably not often important in wine tasting.

Color perception is subject to many influences. In the first place, visual sensitivity to color differences varies considerably from one normal observer to another. (Color blindness is an abnormal type of visual sensitivity.) The individual's inherent perceptivity is modified first by age. Sensitivity to absolute light decreases after 20 years of age. Color perception is constantly modified by learning and conditioning, however; intelligence and experience can compensate for the gradual decline in sensitivity that comes with age. Specialists may, and frequently do, increase their sensitivity by visual training. Another complicating factor is that the preference of the observer for colors varies. Eysench (1941), for 21,060 observers who had been asked to state their order of preference by ranking from 1 to 6, obtained the following average order of preference: blue (1.42), red (2.20), green (3.18), violet (3.92), orange (5.07), and yellow (5.21).

There are also external factors affecting color sensitivity. For example, purity discrimination of the human eye is greatest with a wave length of 570 $m\mu$; it is far less acute immediately below and above this. In addition, the immediate past color experience of the observer is important in achieving constant color evaluation.

The background against which an object is viewed is another factor influencing color perception, for the impression of a color varies with different settings. Color discrimination is thus subject to contrast phenomena—that is, the surrounding color influences our perceptions of dominant wave length and purity. If the background color is the same, the apparent color of the object will be lighter than if the background color is some contrasting color (green on green, for example, looks less green). This is probably of some practical importance, and hence wines of similar color should always be judged against the same background, which should be of a contrasting color.

As will be indicated later, the instrumental approach to color specification is more rational. As yet, however, there is no agreement in the wine industry on the appropriate color for each type or on the acceptable limits in the

luminance, dominant wave length, and purity. Until such standard colors are established, they cannot be specified in terms of brightness, dominant wave length, or saturation. We will, however, suggest some limits later (p. 491). The beer industry, on the other hand, has established such specifications (Stone, 1954).

Odor

The olfactory regions are located in the upper portion of the nasal cavities. During normal breathing, little inspired or expired air gets as high as the olfactory cleft. To force air up into this region, one must sniff. Only by sniffing can one smell attentively and accurately. The olfactory region is yellowish; altogether it covers an area of about 500 sq. mm. It is innervated by the olfactory nerve.

To be odorous, a substance must be volatile, dissolve in the aqueous mucus, and reach the olfactory region in sufficient concentration. There is little correlation between odor and vapor pressure according to Mullins (1955), but there is some parallelism between the lipoid solubility of a substance and olfaction.

There is no classification of odors comparable to the four primary tastes. The Henning (1924) division—spicy, fragrant, resinous, putrid, ethereal, burnt—does not really aid us very much. Crocker and Henderson (1927) devised a system based on eight intensive levels for each of the four primary qualities: fragrant, acid, burnt, and caprylic. The primary utility of this scheme is that each level is defined in terms of a specific chemical. Unknown odors can then be defined in terms of a number consisting of four figures, each of which represents in order the level of intensity found for the four primary qualities. No classification, however, is wholly satisfactory.

The most important aspect of the sense of smell is its extreme sensitivity and selectivity. Our olfactory threshold is about 6,000 to 10,000 times as sensitive as that for taste. The value of 4×10^{-8} mg of ethyl mercaptan (7×10^{-13} mol) per liter of air is one of the lowest recorded. However, we can distinguished smaller differences in tastes than in odors. Odor thresholds and differential odor sensitivity are difficult to determine because it is hard to control the concentration during breathing. Indirect procedures based on sniffing or injection are commonly employed. Mullins (1955) has pointed out that much of the data in the literature is suspect because of doubt as to purity of the chemicals employed.

There are important interactions between odors. Exposure to camphor, for example, raises the threshold for eugenol but does not affect that for benzaldehyde. Odor counteraction based on incompatible pairs of odors is also theoretically possible. Some masking occurs in wine when an intense varietal or a strong off-odor is present. No incompatible odor pairs have been reported in wines, but more evidence is needed.

Adaptation—that is, temporary or permanent loss of sensitivity after exposure to an odor—is a common complicating factor in wine tasting. The rate of adaptation increases as the concentration increases, and recovery is slower than adaptation. Elsberg (1935) showed the duration of fatigue to be proportionate to the length of time the receptors were stimulated. He also noted that if the olfactory regions in both nasal passages were simultaneously fatigued, adaptation persisted longer than if only one were fatigued. During

adaptation there may be qualitative changes in the nature of the odor. Also, when two similar odors are smelled together, one will usually mask the other. However, if the one causing the masking is smelled until adaptation occurs, the other odor can still be identified.

Taste

The taste buds are located primarily on the tongue, occurring on certain raised portions called papillae. The taste buds are about 0.07 mm long and 0.05 mm wide. Each consists of several receptor cells and supporting cells all grouped together in a budlike structure around an inner chamber. The taste bud is innervated with nerve fibers.

While the term "taste" is loosely employed to indicate both tastes and smells (as when we say, "The peach tastes good"), the two senses should be carefully distinguished. If the nose is completely closed, only a limited number of tastes will be perceived. The primary tastes are sweet, sour (or acid), salty, and bitter. There is a variety of evidence in favor of four primary tastes. Inability to discriminate between mixtures of these and other suspected primary tastes is one piece of evidence. Differentiation of the four primary tastes on different parts of the tongue is another. The various effects of narcotics on their detection is still another. Historical and common experience is also in favor of a limited number of primary tastes. Finally, the electrical activity of different fibers is markedly different when stimulated by the primary tastes.

The sensitivity of the taste buds in the various papillae to the different primary tastes varies markedly. Some papillae are sensitive to three or four of the primary tastes, others to only one or two, and some to none. Furthermore, the sensitivity in various parts of the tongue also varies. Sweet and salty tastes are best perceived at the tip of the tongue, bitter at the back of the tongue, and sour on the edges. Thus for the best over-all taste, the whole area must be in contact with the substance or substances being tasted.

The sense of taste, unlike that of sight, is fairly discriminatory. Sweet and sour mixtures can be tasted at the same time. However, if one of the substances is at or near the threshold for discrimination, it will seldom be discerned in the presence of a high concentration of the other. The threshold for a taste substance is defined as the smallest concentration necessary to produce a gustatory response. There are thousands of determinations of thresholds in the literature (see p. 487). Hopkins (1946) reported that there was no correlation between taste acuity and ability to assess palatibility. He attributes this to the fact that olfactory, tactile, and kinesthetic (see p. 486) sensations, plus subjective reactions based on training, prejudice, and other factors, are involved in palatability judgments.

Contrast phenomena have also been reported. For example, when one side of the tongue is stimulated for salt, the other becomes more sensitive to sweet. The opposite is also true. Bitter has been shown to contrast only slightly with the other tastes. Other experiments by Fabian and Blum (1943) seem to indicate that the contrast phenomena depend on concentration. At subthreshold concentrations sodium chloride reduced the sourness of acids and increased the sweetness of sugars. Organic acids had varying effects on sugars. Some acids decreased the sweetness of some sugars and increased that

of others. Other acids had the opposite effect or more. Finally, sugars reduced the salty and sour tastes, but sucrose was more effective than the other sugars in reducing the taste of malic and tartaric acids. In still other cases, adaptation to sugar increased salt, sour, and bitter sensitivity. Adaptation to bitterness, on the other hand, seemed to reduce sensitivity to sugar.

These effects may be of considerable importance with wines where three and sometimes four primary tastes are found. Furthermore, the influence of other substances, such as alcohol, sulfur dioxide, glycerin, etc., on the sensitivity to the primary tastes needs to be considered (p. 506). It should be noted, however, that the methodology for tests of this nature could be improved. Hopkins (1953) for example, did not find any effect of quinine sulfate on the salty taste, but at high salt concentration (1.5 per cent) he obtained reduced bitterness at low quinine concentrations and increased bitterness at high quinine concentrations. The effects noted were not large.

The response of the taste receptors is very rapid, but its speed varies with concentration and the particular primary taste. For sodium chloride it varies from 0.37 to 1.01 seconds, for citric acid from 0.48 to 1.32 seconds. Adaptation (fatigue) is much slower. When flowing solutions are used at low concentrations, the following values are reported for the time required to adapt: salt, 20 seconds; sour, 1.5 minutes; sweet, 1.0 minute; and bitter 1.5 minutes. At higher concentrations fatigue was slower—2.0, 3.0, 5.0, and 2.5 minutes, respectively. Fatigue has been found to be more of a factor in some types of commodity tasting than in others. Beer, for example, seems to have a specific fatigue factor according to Lane *et al.* (1954). There is no doubt that fatigue is of critical importance in wine tasting—particularly in dessert wines where the sugar and alcohol contents are high. Some people find red wines particularly fatiguing because of their high tannin content. Others find sparkling wines tiring because of the carbon dioxide. Finally, the relatively high sulfur dioxide content of white table wines makes it difficult to judge more than a few at a time.

The effect of the temperature of the food or beverage on the primary tastes is surprisingly variable. The optimum for sucrose and hydrochloric acid is reported by Cameron (1947) to be as high as 35° to 50° C, while that for saltiness varies from 18 to 35° C and that of quinine is about 10° C. Raising the temperature in the normal range of food temperatures appears to increase the sweetness of a solution but to decrease its saltiness and bitterness. Temperature also influences the rate of adaptation: with sugar, higher temperatures reduce the absolute threshold and the time for adaptation.

The Common Chemical Sense

The mucous membranes of the mouth, nose, eyes, etc., are responsive to a number of irritants. Free nerve endings of the trigeminal nerve appear to be stimulated, causing sneezing, cooling, choking, pain, watering of the eyes. This sense differs from taste in that there are no specialized nerve endings. It is easily demonstrated as being different from smell by using a menthol nose inhalator. On inhalation, there is first a cooling sensation (common chemical effect) and later a menthol odor (olfactory sense). If the nose is tightly closed and the inhalator placed in the mouth and air drawn through it, only the cooling effect will be noted. If the olfactory nerve is cut, the common chemical sense still remains.

Taste buds are reported to respond to 3 *M* ethyl alcohol (Parker, 1922). For irritant action of the common chemical sense a 5 to 10 *M* concentration is needed. Besides ethyl alcohol, the following substances present in wines or brandies appear to affect the trigeminal nerve: sulfur dioxide, acetaldehyde and other aldehydes, and some of the higher alcohols.

Other Senses

Kinesthetic. The nerve endings in the muscles, tendons, and joints also respond to foods. Those located in the muscles of the tongue and the joints of the jaw may be concerned with wines. The toughness of meat and the crispness of lettuce are examples of responses of this sense. It is difficult to study because the nerves cannot be removed. However, cocaine is known to prevent the response. The high extract content of very sweet wines may influence this sense.

Tactile. Nerves located in the mucous membranes of the mouth and nose are receptive to tactile, or touch, sensations. Our response to smoothness, creaminess, and oiliness is believed to be cutaneous in origin. Here again the high extract content of sweet wines may give a response, or the thinness of very low alcohol wines may also be detected. The tang of carbonated or sparkling wines is probably indicated through this sense.

Response to Temperature. Receptors sensitive to temperature differences are located on many parts of the body, particularly in the mouth. Temperature is important in wine tasting in several ways—first, for its own effect. We prefer, owing to experience, that certain wines be cold, others warmer. Second, temperature influences the volatility of the odorous substances in the wines, and to be smelled a compound must be volatilized. Third, the sensitivity to the primary tastes varies markedly with temperature (p. 485).

Physiological and Psychological Factors Affecting Sensory Perception

Sensitivity to tastes and odors varies markedly among individuals. In addition to the individual's inherent capacity, there are many external factors influencing his sensitivity. Environmental influences—food preferences prevailing in the region, which are dictated largely by the foods available, as well as the more specialized preferences of the family group—will to some extent affect sensitivity. Inherent capacity is modified, and personal preference determined, by environment, experience, and training. It may be developed to a high degree of acuity; it may remain undeveloped through lack of interest or opportunity; or it may suffer damage from disease or accident. While the various major factors affecting taste and odor sensitivity are discussed separately, it will be obvious that they are intricately interrelated. Age is closely involved with experience and training, with ability to concentrate and remember. In addition to the modifications that one's sensitivity may undergo, inherent ability to concentrate and remember can also be modified by experience, motivation, and training, with the whole physiological and psychological complex changing as the individual ages.

The individual's highly subjective opinion as to the desirability (quality) of a certain wine should not be confused with the more or less objective assessment of panels of specialized judges. The above factors influence both

groups, but the discussion here is how to minimize their interference (or how to use them for greater efficiency) with the second group.

Individual thresholds. Often, one of the first tests performed in selecting panels for critical sensory experiments is to determine the thresholds of the prospective panel members for the primary tastes. Odor tests are not usually performed because the methodology is more complicated. Tasters of low specific sensitivity are usually dropped from the panel.

The averages given by Hopkins (1946) show the ranges of the thresholds to be: sucrose, 0.0192 to 0.0195 *M;* caffeine, 0.0008 to 0.0018 *M;* glutamic acid, 0.0008 to 0.0010 *M;* tartaric acid, 0.00020 to 0.00026 *M;* and sodium chloride, 0.0192 to 0.0199 *M*. He also showed a significant correlation to exist between an individual's sensitivity to sour, salty, and sweet substances. Sensitivity to bitter was only slightly associated with sour and not at all with salty or sweet. In his tests, females were better sour tasters than males. Usually no differences between males and females are obtained.

An example of group variability was given by Rouleau (1953), who studied the responses of a family of good tasters and one of less discriminating tasters. In a test using paired samples of beer, the former family scored 58 per cent correct responses, the latter 36 per cent.

Age. Differences with age have been demonstrated, but to what extent they are due to age and to what extent to experience is not known. There seems to be an improvement in sensitivity up to about 25 or 30 years and a decrease after 50 or 60. It is certain that there is a decrease in threshold—that is, an increase in sensitivity—with practice. Probably the older taster learns to recognize smaller amounts of the substances by concentrating on looking for them. Rouleau (1953) preferred beer tasters to be between 35 and 45 years old, but he had one good 74-year-old taster. Tilgner (1957) has found older (over 45) tasters to be *more* sensitive to sugar than younger (20 to 35) tasters. However, he also reported them less sensitive to sour and bitter.

Experience. The importance of experience, which most experts are agreed upon, is difficult to measure. It is common belief, for example, that inexperienced tasters prefer wines of a pronounced sweetness and low astringency. According to Puls (1939), German tasters prefer wines of high total acidity, wines that French tasters would find disagreeable. In certain parts of southern Italy and Catalonia, tasters prefer wines of high alcohol (15 to 16 per cent) as a table beverage.

Moreover, these differences are constantly changing. Puls (1939) reported that the German taste was no longer for wines that have been aged for several years in the cask before bottling, but rather for early-bottled, young, light-colored fresh wines. Recent personal experience of one of the authors (M.A.A.) with German wines seems to confirm this trend.

It has been shown that there are differences between wine judges and tasters who are not wine judges (Filipello, 1957). These studies are suggestive and need to be extended. Wagner (1950) says that tasters must be given systematic training if reliable results are to be obtained. Brunet (n.d.) has expressed the opinion that judges who are accustomed to tasting high-quality wines are apt to underrate ordinary ones and that those who are accustomed to ordinary wines will overrate fine ones. He also believes that those who

are accustomed to rich, heavy-bodied wines may not perceive the fine points of lighter wines. See also the discussion on the flavor profile in Anon. (1958).

Experience and training are known to be extremely important in the effectiveness of certain kinds of panels. Engel (1933) noted that in the tasting of five wines the discriminatory ability of the judge could be much improved by training.

It is quite possible for the men in the winery to taste their own wine so often and to the exclusion of all others that they become accustomed to slowly developing off-characteristics. One of the authors (M.A.A.) observed this in the proprietor of a sparkling-wine cellar, who was proud of the special taste of his wine, which he maintained by using his own cultures. The fact was that his wines were all altered by *Lactobacillus* activity and were spoiled because he was using a contaminated culture. Lefebore in Rouleau (1953) cites an example of a brewery where all the technicians were insensitive to a spoiled character in their beer—in fact, could not even be persuaded that it was present as a spoiled characteristic.

Memory. Memory is always listed as one of the important requisites of a taster (Mathieu, 1911; Morrow, 1936; Got, 1953). To identify an odor or a taste, we must be able to remember it from previous experience. Experience and training are necessary. Memory also enables us to differentiate more degrees of quality; this, again, depends upon experience. Finally, in many comparative types of judging—that is, between two or more samples—memory is a great aid to accurate ranking. However, Filipello (1957) has shown that even experienced tasters may lack ability to grade quality accurately when only one sample is available. The great advantage of a keen memory for the more obvious odors and tastes is that it saves the taster time for consideration of the less easily recognized characteristics. Tasting and smelling are fatiguing; hence, anything that shortens the time involved in recognition and evaluation should help maintain sensitivity.

Concentration. There is no doubt that considerable concentration of attention is necessary when small differences are to be detected. Elsberg (1937) showed that for odor at least, the threshold was lower when the observer concentrated on "thinking" of the test odor. Since these tests were done with his blast-injection technique and blank tests were interposed, the results are highly suggestive. If they are reliable, they strengthen the suggestion that the wine taster should examine the sample or samples methodically, concentrating on the examination of a single characteristic at a time.

Motivation. There is more to judging wines than simply accurate use of the senses. The psyche also plays a part. The taster who does not care or want to find differences in character or quality will not find them or will differentiate poorly between wines of varying character or quality. Experience and motivation are necessary to counteract this. Companies planning to use panels for difference testing should provide for rewards in order to secure the optimum results. Competition between tasters is also effective in improving and maintaining performance. Extra pay, time off from regular work, and recognition of tasting ability are also useful stimulants to greater efficiency. It is absolutely essential, for consistent results, that the tasters be interested. As Beattie (1949) so aptly puts it, "The bored are more to be feared than the 'blind' [taste-blind]."

Time of day. Goetzl (1949), using the Elsberg technique, reported that sensitivity to odor was lowest immediately after meals and increased to a maximum just before eating. Whatever may be the final conclusion regarding the significance of his results on the relation between time of eating and olfactory acuity, there can be no doubt that this period is a variable which must be controlled. All the older enologists, such as Grazzi-Soncini (1892) and Brunet (n.d.), caution against eating strongly spiced foods before tasting, and Grazzi-Soncini says no foods, not even bread, should be eaten during tasting. Rouleau (1953), for beer, cautions against eating during tasting, recommends eating very little for lunch if it is necessary to taste afterwards, and warns against eating highly spiced foods. Mathieu (1911) preferred tasting in the morning. Rouleau (1953), too, prefers the morning hours, feeling that the tasters are more likely to be relaxed and in good mental condition. In our experience, the period just before lunch is the most desirable time of day for tasting.

Use of tobacco. Hopkins (1946) found no taste differences between smokers and nonsmokers. Grazzi-Soncini (1892), however, said categorically that smoking before or during the tasting reduces the value of a judge's results, although he gave no data. Rouleau (1953) did not find that smoking had any influence on beer tasters, but he did show that the sensitivity of nonsmokers was markedly reduced when other tasters smoked in their presence.

Probably the nonsmoker would be thrown off either by smoking or by smelling tobacco smoke. The regular smoker probably compensates for the tobacco smell and suffers little harm. Furthermore, the psychological effect of temporarily depriving a smoker of tobacco may make it difficult to measure any effect of smoking *per se*.

Loss of Sensitivity. Disease or accident may result in ageusia (loss of taste), hypogeusia (diminished taste), or parageusia (altered taste). The most important and well-documented case of specific taste blindness is that for the bitter taste of phenylthiourea. This has been shown to be a recessive genetic character. About one third of the population is only slightly sensitive to this compound. However, there is no evidence that such individuals cannot taste bitter compounds in wine, although individuals do differ markedly in their expressed reactions to the apparent bitter taste of some red wines. Psychological factors, such as anger or sadness, and physiological factors, such as fatigue or a cold, cause an immediate decrease in sensitivity.

III. THE CHEMICAL COMPONENTS OF WINE: THEIR ORGANOLEPTIC DETECTION AND DESCRIPTION

In this section an attempt has been made to bring together the available information on the components influencing the color, odor, and taste of wines. In connection with each, the concentrations normally present, as well as those necessary for organoleptic differentiation, are given. Where possible, the terms used to describe the sensation produced are stated.

Descriptive terms have always been criticized as being one of the weaker points in organoleptic tests. Terms such as *pleasant*, *agreeable*, or *fine* lack precision and give us little real information. At times such terms are undoubtedly employed because of uncertainty as to the organoleptic impression.

In other cases such words may have a precise meaning for one taster but not be easily comprehended in the same context by others.

In the European literature much emphasis is placed on the use of descriptive terms for the various characteristics of wines. However, even in European glossaries of tasting terms, we find many overlapping or vague terms. See, for example, Brunet (n.d.), Got (1953), Jungendfein (1937, Klenk (1950), Mensio (1957), Office International du Vin (1959), Puls (1939), Troost and Wanner (1955), and Ventre (1931). Wagner (1950), who has studied this problem for the German language, rejected such ambiguous and imprecise terms as *clean, indifferent, insipid, pleasant* and *unpleasant, pure* and *impure, fine, delicate, mild, charming, exciting,* and *feeble.*

Terms used in the organoleptic examination of foods should have only a single connotation. We have, therefore, used only terms that are clearly necessary. Investigators with greater organoleptic skill and literary expression will undoubtedly develop words for finer gradations of meaning.

General terms that have some sense of form may occasionally be useful: *full, full-bodied, thin, round, rich, smooth, sandy, sharp,* etc. Touch, temperature, and pain reactions can be utilized if clearly understood: *watery, soft, hard, astringent, harsh, gassy, pungent,* etc. But the most useful terms are those that have specific taste and odor meanings: *acid, salty, sweet, dry, rough, goaty, mousy, aldehyde, fusel oil, sulfide, moldy, yeasty, fruitlike, winey, spicy, resin, caramel, earthy, smoky, lemon, vanilla, woody, corked, varietal-odor, lees-odor, stem-odor, pomace-odor, frost-odor, raisin,* etc.

Although descriptive English words for wines are lacking because the English language developed in a non-wine-producing area, there is no reason why the necessary terms cannot be developed. Successful use of such terms, however, depends on each having a clear and unequivocal meaning. This may prove difficult in California whose vintners represent diversified cultural backgrounds in which similar terms are defined differently. The present definitions have therefore been made with particular emphasis on clarity of definition and on the avoidance of two or more words describing the same sensory quality.

Color

The color of a wine is one of its most important attributes. Certain deviations from the normal for the type are tolerated, but large deviations are rejected. Thus, a white table wine that is dark brown is seldom acceptable, even by the least-experienced wine consumer. This is natural because extreme variations in color have been found to be associated with undesirable changes in flavor or odor or both.

The human eye is a most discriminating organ, capable of distinguishing a vast number of different shades of color, according to Mackinney and Chichester (1954). However, the eye is not a very good quantitative instrument for permanent records. The modern tendency is to specify color in the physical terms of dominant wave length, purity, and luminance (or lightness). The suggested standards are only tentative, but it is hoped that they will stimulate the interest of industry in establishing standards and acceptable limits.

Naturally, the color-blind person cannot be expected to differentiate a wide range of colors, though he may be able to differentiate certain ones,

the exact shades of which escape him, on a scale of grayness. Now that rapid methods for describing color in terms of its fundamental physical attributes are more easily available, the instrumental approach is more rational.

Whites. All white wines contain some yellow flavonol pigment (quercetin or isoquercetin), but in some the concentration is low. Others are a full yellow or even a gold color, and in still others increasing amounts of some brown pigment seem to be mixed with the yellow. In some cases oxidation products of flavonol pigments are present.

TABLE 1

APPROXIMATE COLOR CHARACTERISTICS*

Observed color	Lightness, per cent	Dominant wave length, mμ	Purity, per cent	Duboscq color†	Klett reading‡	Munsell§	
						Hue	Value/chroma
Light straw yellow	93	574	12	8	63	5.0Y	8/6
Medium yellow	91	575	15	14	120	5.0Y	8/10
Light gold	89	576	20	18	200	2.5Y	8/12
Medium gold	80	576	25	25	275	2.5Y	7/10
Light amber	85	578	30	20	220	5.0YR	7/10
Medium amber	80	579	40	40	300	7.5YR	5/8
Dark amber	50	580	70	100	450	7.5YR	5/6
Medium pink	93	594	3	70	150	7.5R	5/10
Light red	80	630	5	100	200	2.5R	4/8
Medium red	60	610	15	190	325	5.0R	3/6
Tawny	75	585	15	100	250	7.5R	3/10

Source of data: Amerine *et al.*, 1959.
* The transmission curves for the white wines were made in a 20-mm cell or calculated to 20 mm. The reds were in a 1-mm cell. It should be stressed that these are approximate values based on a limited sampling.
† Using the white or red color standards of Winkler and Amerine (1938).
‡ Using a 42 filter for the whites and a 46 filter for the reds.
§ In the Munsell system the first figure and letter represent the hue (roughly dominant wave length), the second figure is the value for luminance (the brightness), and the third the chroma (purity or saturation).

Light straw yellow. The lightest shade of our very lightest-colored wines, such as chablis and White Riesling. Approximate color characteristics are given in table 1 for this and later colors.

Medium yellow. Still a rather light-colored wine. Most of our California white table wines fall in this group, particularly dry sauterne, Semillon, and Sauvignon blanc.

Light gold. A fuller yellow, characteristic of the sweeter California table wines, such as sweet Sauvignon blanc, sweet Semillon, or chateau-type sauterne.

Medium gold. Full gold color, found in some of our domestic and in many imported sweet table wines. Some muscatels likewise have this shade.

Light amber. Most of our lightest California sherries are of this color—a medium yellow with some admixture of brown.

Medium amber. Many commercial California sherry and muscatel wines have a medium amber tint.

Dark amber. This is an uncommon shade in California wines except for sweet vermouths, which also have a slight reddish tint. This is undesirable if too pronounced. Marsalas are dark amber.

Once in a great while a young white wine may have a greenish tint, due

to residual chlorophyll. In most, if not all, cases this is a desirable characteristic. Because of California's warm climatic conditions such a tint is rare here, but it is found in a few German wines.

Reds. The red color is due to a mixture of anthocyanin pigments, primarily malvidin-3,5-diglucoside, malvidin-3-monoglucoside, and 3-delphinidin-3-monoglucoside, as well as their anthocyanidins, malvidin and delphinidin. Because of their structure, the color of these pigments varies with the pH of the solution. Their isoelectric point is about 5.2 at which pH they change from a red to a blue, bluish-green, or grey color. The color is also affected by metallic ions, the amounts and relative proportions of the different pigments, the amount of sulfur dioxide present, and other variables. The age of the wine is also a factor. Young red wines are often violet-red. Within a few months they lose the violet tint and become full red. Still later, particularly if oxygen is allowed to come in contact with the wine for long periods, a brownish or amber tint is noticed. Very old red table wines that have been aged in the bottle may deposit most of their red color and retain only a pink-amber hue. Red dessert wines aged in warm cellars may deposit all of their red color and appear as a light amber or later even a dark amber.

Since *rosé*, or pink, wines are becoming increasingly important in California, greater attention should be paid to their color characteristics. This is particularly necessary to prevent blends of red and white wines from being sold as *rosé* wines. It may not be too difficult to distinguish these by accurate measurements of their color characteristics. Wines with a pale, faded pink color are almost always such blends. It is difficult, however, to set stringent limits on the color of pink wines because of the variation in color characteristics of the different varieties employed. Thus, Petite Sirah, Pinot noir, and Grenache have a distinct orange tint, while Gamay, Cabernet Sauvignon, Mondeuse, Carignane, and Zinfandel (if not very clearly overripe grapes) produce wines without this orange hue. Furthermore, the same variety grown in a warm climate or picked late in the season will have more of an orange or brown hue than when grown in a cool region. At present, *rosé* wines can best be classified as follows:

Low pink. A distinct pink but with an intensity below the acceptability as a *rosé.*

Medium pink. Most genuine commercial samples of the *rosé* type fall in this group. Orange is a modifying tint of certain varieties. Brown should never be an attribute of this type of wine.

For red wines, the following scale may be applied:

Low red. The depth of color is above that accepted for *rosé* wine but lighter than most of the standard types of red wines.

Medium red. The depth of color of a standard red wine, similar to a California burgundy, claret, or Zinfandel.

High red. This color is characteristic of blends and very young wines. They frequently have a blue or purple tint, which may be due to being very young, to the use of varieties with purple tints (such as Alicante Bouschet, Alicante Henri Ganzin, Almission, Grand noir, Salvador, and occasionally Petite Sirah), or sometimes to the very high pH of the wine. Wines that have undergone an extreme malo-lactic fermentation may have a pH of 3.9 to 4.4 and a purple hue. The dominant wave length may then have to be given in

terms of its complementary, i.e., 504c, or, if lacking purple, may be 640 to 655 mμ. These colors are often equivalent to Munsell 2.5 R 3/10.

Old wines, table or dessert, may be amber or brownish. This is often called a "tawny" color in dessert wines and an "onion-peel" or "eye-of-the-partridge" color in red table wines. The color can be induced by baking the wine, by using certain grape varieties of low color, and by excessive aeration, as well as by aging. Whether or not the color characteristics (dominant wave length, etc.) would be the same for a naturally aged wine and a baked one is not known. For approximate color characteristics see table 1. For further information see Amerine *et al.* (1959) and Crawford *et al.* (1958).

Appearance

In wine evaluation the term "appearance" is thought of as a quality factor separate from color. It refers to the degree of clarity of a wine, or its freedom from suspended or precipitated material. This aspect of wine quality is very important for the general consumer, who has by long experience discovered that a cloudy wine is usually a defective wine. There are, however, exceptions, and the less sophisticated wine drinker is likely to condemn a wine with a deposit without even tasting it.

Suspended material in commercial wines is, almost without exception, an indication that the wine has been improperly finished or is spoiled or spoiling. Cloudiness due to the unfinished nature of a wine may be caused by failure to remove tartrates, yeast cells, bacteria, or amorphous material that has become insoluble at the higher alcohol content of the wine. Cloudiness due to spoilage organisms results from the population of the microorganisms themselves or to their metabolic by-products. There are also types of cloudiness that result from excess metal content or enzymatic reactions.

Deposits are quite another matter. A sediment *may* arise from unfinished or spoiled wines. Deposits may also arise in red wines from the natural precipitation of their anthocyanin pigments. In white wines the deposits may consist of quite harmless tartrates. In either case the deposits do not influence the quality of the wine *unless* (and this is important) they are disturbed and are suspended in the wine while it is being consumed. This can usually be avoided by proper decanting. However, we must recognize that the usual wine drinker does not have the knowledge or time to do this properly. With white wines, where such deposits are more easily preventable, the deposit does constitute a negative quality factor. With red wines, on the other hand, where deposits during bottle aging are unavoidable, they should not be so considered *when the wine has been aged*. This obviously does not apply to many commercial wines in California.

Wine tasters can gain much useful information by inspecting the appearance of a wine. Various types of wine spoilage result in typical kinds of cloudiness. Rod-shaped lactobacilli usually give a silky, shimmering appearance. In extreme cases the wine may become oily and viscous. Other lactobacilli cause cottony deposits. Yeast growth in white wines gives a milky appearance, often with a granular deposit. Excess ferric phosphate also results in milky-white wines. A similar haze develops with excess copper; however, a reddish-brown precipitate is also frequently present. Cloudiness due to precipitation of potassium acid tartrate and calcium tartrate may also be milky in appearance, but the deposit is crystalline.

The terms used to describe the appearance of a wine are as follows: *brilliant,* the wine is free of all suspended material; *clear,* a very slight haziness may be seen if the wine is closely observed; *dull,* the haziness is easily seen; *cloudy,* there is so much suspended material that some is found as a deposit. For a given wine it is frequently desirable to determine when a change in appearance begins to occur during filtration. This was formerly done by taking successive samples and inspecting them visually. Nowadays wineries use photoelectric colorimeters to measure the per cent transmission, or nephelometers which measure reflected light. As soon as the per cent transmission begins to decrease, or the per cent reflected light to increase, one knows that the efficiency of the filtration has decreased.

Odors—Undesirable

The odor of a wine should be free of any foreign or unnatural scent. A wine's odor is derived from constituents of the variety of grapes from which it was produced, from compounds developed or added during fermentation, and from materials produced by aging. The definition of what is natural is not free of ambiguity. A woody odor is commonly accepted and even desired in sherries and certain red wines, but in white table wines it is not. The odor of brandy constituents is undesirable in still white table wines but is found in many sweet sparkling wines. The terms listed and described below differ more or less from those in other works, even in English. This is not strange, since many definitions for the same word have been employed by various authorities. Grazzi-Soncini (1892) wrote that the confusion was already great in his day, and it has certainly not diminished with time. We have tried to define each term clearly to reduce the confusion as much as possible.

The common off-odors found in wines are (1) sulfur dioxide, (2) hydrogen sulfide, (3) mercaptan, (4) earthy, (5) mousy, (6) oxidized, (7) cooked, (8) raisin, (9) stemmy, (10) woody, (11) green, (12) yeasty, (13) lees, (14) moldy, (15) stagnant, (16) acetic, (17) butyric, (18) lactic, (19) pomace, (20) corked, (21) fusel, (22) rubbery, (23) sophisticated, (24) filter-pad, (25) sack, (26) charcoal, and (27) bitter almond.

Sulfur dioxide. Nearly all white table wines, many red table wines, and even some dry and sweet dessert wines contain sulfur dioxide. Little of the sulfur dioxide is present as the dissolved gas but is found as sulfite ($SO_3^=$) or bisulfite (HSO_3^-) ions, called free sulfur dioxide, or is fixed to dextrose or to acetaldehyde (as α-hydroxyalcoyl sulfuric acid) and is then called fixed sulfur dioxide. The relative amounts present as free and fixed depend on the composition of the wine, the amount added, and the length of time since it was added. It is the free sulfur dioxide which is mainly responsible for the objectionable odor. Since sulfur dioxide also affects the common chemical sense, there is a tendency to sneeze when examining wines high in free sulfur dioxide.

Berg *et al.* (1955*a*) reported a threshold concentration of 0.0011 gm per 100 ml of sulfurous acid in water, or 8.6 ppm as sulfur dioxide. The water had a pH of 5.8 to 6.0. Using a water solution of 0.002 gm per 100 ml of sulfurous acid, they found a difference threshold of 0.002 gm. The per cent free sulfur dioxide must have been very high in these solutions. However, it is difficult to maintain the concentration constant at such low levels. In

wines, on the other hand, the difference threshold was much larger. Using a white wine with a total sulfur dioxide content of 52 ppm and a free content of 14 ppm, they reported a difference of 207 ppm in total and 120 in free as the minimum detectable. Using a red table wine with 32 ppm total and 2 ppm free, they found the minimum differences to be 100 ppm total and 58 ppm free. These differences are considerably higher than those of many experienced tasters. More data would be desirable. Gentilini (1949), for example, found 20, 50, and 100 ppm total sulfur dioxide distinguishable in dry white wines and 50, 100, and 200 in dry red wines. Berg *et al.* (1955*a*) noted that experienced tasters were no better than inexperienced tasters in detecting sulfur dioxide. In their tests, however, a triangular taste test was employed, and fatigue may have been a limiting factor. Boggs and Ward (1950) demonstrated the effect of fatigue to be important in tests with potatoes containing sulfur dioxide.

Hydrogen sulfide. Reducing conditions during alcoholic fermentation permit formation of hydrogen sulfide. When elemental sulfur is hand-distributed in the vineyard to control powdery mildew, it is often applied unevenly and excessive amounts get into the fermenters. In high-acid musts, especially, large amounts of sulfide are formed. If the wine is aerated promptly and separated from the lees and if sulfur dioxide is added, the odor is usually dissipated. Amerine (1958) estimated at less than 1 ppm the hydrogen sulfide content of wines with only slightly detectable sulfide odor.

Mercaptan. In a few cases ethyl sulfide is formed. This has a persistent and very unpleasant odor which is difficult to remove. It is often described as a garlic-like odor. To distinguish between hydrogen sulfide and mercaptan is often difficult. The following chemical filter procedure was suggested by Brenner *et al.* (1955). Place about 40 ml of wine in each of three glasses; to the first add 5 ml of water, to the second 5 ml of a 5 per cent solution of copper sulfate, and to the third 5 ml of a 5 per cent solution of cadmium sulfate. Stir. If any difference in odor is noted between the check and the glasses to which copper or cadmium sulfate has been added, one may suspect hydrogen sulfide or mercaptan or both. Now smell only the latter two glasses. If there is no difference between them, mercaptan is probably responsible for the odor. If there is a difference, both hydrogen sulfide and mercaptan are probably present. To remove them, aeration is the first step. In persistent cases activated charcoal may be used. Amerine (1958) found less than 0.05 ppm of mercaptan in wines that had a slight off-odor. This agrees with recent work with beer where only 0.03 ppm was reported.

Earthy. While this is often called an earthy taste (*goût de terroir, Erdgeschmach*), it is a true odor. The cause is not known. Peglion (1900) suggested earth on the grapes or growth of *Cladothrix odorifera* in the casks. Moncrieff (1951) suggests that the earthy odor may be due to *Actinomycetes* sp. Brunet (1941) reviewed the various opinions as to its origin in France. He was inclined to think it originated in the soil but noted that old tartrate incrustations and sulfuring had been suggested. In our experiments, wines from Davis had been most likely to have an earthy odor, though some Napa Valley wines have it. Washing the grapes reduces, but does not completely remove, the odor. This odor is perceived only when the wine is warmed in the mouth, and the sensation is *apparently* perceived in the roof of the mouth

after the wine is tasted. This is why it is sometimes considered a taste rather than an odor. Jungendfein (1937) correctly identifies this odor as one characteristic of the region. He suggests yeast strain, cellar treatment, and soil type as possible sources. The last seems the most likely cause.

Mousy. The odor of mouse's urine is associated with certain types of spoilage by *Lactobacilli*. The odor is often best discerned by rubbing a little of the wine on the hands where the heat of the rubbing releases the characteristic odor. The odor has been ascribed to a variety of chemical compounds including formaldehyde, acetamide, acrolein, etc. Zaslavskiĭ (1955) favored acetamide as one cause but indicated that excessive oxidation was also a factor. He reported the odor to have been produced experimentally by adding peroxide to low-acid grapes (produced by adding alkali) that had been fermented at high temperatures.

Oxidized. Formation of acetaldehyde due to exposure of wines to oxygen or to action of certain film organisms leads to this odor. Where the exposure is temporary, the wine "recovers." Prolonged exposure leads to greater accumulation of acetaldehyde and to the formation of secondary products; white wines darken in color and red wines become brown. There is some resemblance between an oxidized odor and that of prolonged aging in the wood in contact with air. The distinction between the two odors can be established only by experience and by knowledge of the appropriateness for the type. In white table wines an oxidized odor is always a defect. In sherries and similar types it is, in part, a desirable characteristic. Temporary oxidation after bottling is often called bottle sickness. Old white wines that are already darkening in color are said to be "madeirized." This term is inappropriate, however, because modern Madeiras are made primarily by baking, not by long aging in the cask. The extreme oxidation odor of sweet wines produced by long aging in the cask is called "rancio." This is undesirable in white wines, but certain red sweet table wines, such as very old Banyuls and Priorato, seem to require a rancio odor.

Cooked. Wines which have been baked, or to which reduced must or concentrate has been added, have a cooked or baked odor. It is not too difficult to distinguish this from a raisin odor, though both have a caramel "tone." Neither is it the same as the overheated pomace odor of red wines which have been allowed to get too warm during fermentation (see p. 498).

Raisin. This occurs in wines made from raisins or from overripe grapes. In hot years, when susceptible varieties raisin on the vine, the odor is common. Caramel-like compounds are apparently responsible for the odor.

Stemmy. This odor has not been associated with any chemical compound. Berg and Webb (1955) suggest that it is a "modified complex of the bitter and green sensations." Closing the nose prevents its discernment, which must mean that it is due to olfaction not gustation. Stemminess is uncommon in California wines. In some cases at least, it appears to be due to long contact of the wine with the green stems. On the other hand, there are wines which have been so fermented and which do not have this odor.

Woody. This odor—as either a desirable *or* undesirable characteristic—seldom occurs in California wines because most of our commercial wines are aged in cement, metal, or large redwood containers. We know very little about the odor of oak, but vanillin does appear to be a component of the typical

odor of new oak. For red table wines intended for bottle aging, it has frequently been suggested that a trace of this odor is desirable. Sherries also apparently profit by a trace of oak odor. In white table wines it appears most often as a defect. In the tawny types of ports a trace seems to be desirable. In ruby ports its desirability is less certain.

Green. This odor is not commonly found in California wines, although in wines of the 1948 and 1955 vintages it was sometimes observed. It is undoubtedly associated with wines made from immature grapes. Berg and Webb (1955) suggest that the smell is that of leaf alcohol and leaf aldehyde. It is frequently noted in German, Austrian and other northern European wines of the cooler seasons.

Yeasty. Fermenting wines have a very pronounced odor of yeast, which tends to disappear after one or two rackings. During fermentation in the bottle or tank sparkling wines are also yeasty, but they, too, lose this odor after they have been disgorged or after the lees have been filtered off. Less common is a yeasty odor due to undesired growth of yeasts in finished bottled wines. These odors are often described as *fermentation* odors.

Lees. If the yeast is allowed to remain too long in contact with the wine, two conditions may arise. Most commonly a hydrogen sulfide or mercaptan odor will develop. Sometimes, however, the odor cannot be associated with hydrogen sulfide or mercaptans but can only be described as a pungent, or "lees," odor. Both may be present at the same time to make the identification more difficult. An unpleasant bitter aftertaste is sometimes associated with the lees odor.

Moldy. The most common cause of a moldy odor in a wine comes from placing the wine in a container where *Penicillium* is growing. Occasionally this and other molds may be present on the grapes in sufficient amounts to give a moldy odor to the wine. It is very difficult to remove this odor from wines.

Stagnant. Water-logged tanks, if not thoroughly scraped, will taint wines placed in them. Since growth of slime bacteria in water-logged tanks only occurs in cases of very careless operation, it is seldom encountered.

Acetic. The acescent odor of ethyl acetate is familiar to almost all wine tasters. Its vinegar-like odor makes it easy to identify. In winery practice the amount of acetic acid present is commonly measured. Since acetic acid and ethyl acetate are formed simultaneously by *Acetobacter* and since they tend to establish an equilibrium in the wine, the amount of acetic acid present is usually taken as a measure of spoilage. In California the present state limits are 0.110 gm per 100 ml for white table and dessert wines and 0.120 gm for red table wines; the respective federal limits are 0.120 and 0.140 gm. Experienced tasters are sensitive to much lower amounts than this—often to as little as 0.060 gm. An ethyl acetate limit of 150 to 200 mg per liter has been suggested. However, Berg *et al.* (1955*a*) reported a threshold of 8 to 11 mg per liter of ethyl acetate in water. On the other hand, Hinreiner *et al.* (1955*a*) found a difference threshold of 175 to 200 mg per liter in wine.

Butyric. The odor of rancid butter is seldom observed in wines, but occasionally a butyric or butyrate ester odor is found. Sweet table wines seem most prone to this defect, particularly if they are low in sulfur dioxide and unpasteurized.

Lactic. There appears to be a distinction between a lactic taste and a lactic odor, which is sometimes described as a sauerkraut or goaty odor. Lactic acid is not very volatile, though it has an appreciable odor, which is possibly derived from secondary odorous by-products. High-lactic wines are often slightly sweet.

Pomace. It may be possible to classify odors from the pomace into those derived from too long a period on the pomace and those resulting from too high a fermentation temperature in contact with the pomance. Since they usually occur together, they are classed as one odor here. Excessive bitterness associated with these odors clearly belongs under taste (p. 507).

Corked. About one in a thousand bottles of wine aged in corked bottles for more than two or three years will have a corked taste. *Penicillium* sp. and *Aspergillus* sp. in the cork seem to be associated with this taste. Moreover, this odor is usually found in bottles that have had excessive ullage.[3] This suggests that insect infestation and physical defects in the cork are partially responsible. Placing some of the wine between the palms of the hand and rubbing quickly reveals the characteristic odor.

Fusel. Normally this odor is associated with dessert wines fortified with spirits that are high in higher alcohols. That this is altogether an undesirable odor is questionable. Guymon and Heitz (1952) reported a case of a red table wine that contained a detectable amount of higher alcohols. Many Portuguese port wines have a distinguishable fusel-oil odor, and it is apparently not a negative quality factor. Changes that occur during aging may ameliorate the fusel-oil odor to make it a positive quality factor. However, in most California table and dessert wines a detectable odor of higher alcohols is a negative quality factor.

Rubbery. This odor was first described in California wines by Brown (1950), but the expression had been used by wine makers since at least 1947. He stated that it was associated with certain varieties of grapes. Berg and Webb (1955) associate the odor with dessert wines of very high pH (over 4.1), and this seems to be correct. There is probably some oxidation and subsequent polymerization of the oxidation by-products. Occasionally table wines with a pH of 3.9–4.2 have an odor reminiscent of rubber.

Sophisticated. Few wines today can truly be called sophisticated. However, the addition of some constituent which is foreign to the wine can properly be called sophistication or adulteration. Berg and Webb (1955) suggest that weeds harvested with the grapes may give this odor. Eucalyptus trees near a vineyard are another example. Table wines with a vermouth odor (owing to storage in vermouth casks) are still another. Wines of *Vitis vinifera* stored in casks that have contained Concord, fruit, or berry wines and sold as *vinifera* wines could properly be called "sophisticated." Actually, adulteration is extremely rare in California wines, though sparkling burgundy flavored with imitation fruit essence was detected in the period before World War II. Use of extract of vanilla or of vanillin in cream sherries is also an undesirable practice which can be easily detected. Adding coriander seeds or dried elderberry flowers to sweet medium-alcohol wines in southern France to imitate a muscat aroma was reported by Cordonnier (1955, 1956). These could be detected by paper chromatography.

[3] Ullage is the loss of liquid from barrels, casks, or bottles.

Filter-pad. Wines filtered through unwashed asbestos or paper filter pads often have a special odor and a chalky taste. They may also be slightly oxidized, but the filter-pad odor is very characteristic and easily recognizable. A similar odor of wines filtered through cellulose or cloth pads is less common but no less objectionable. It is better to keep the first wine through the filter separate or to recirculate it into the tank before bottling starts.

Sack. A number of our experimental wines have been described as having a "sack" odor—reminiscent of a burlap sack. Since the same odor has been assigned to the same wine on several tastings, without the taster knowing the previous results on the wine, we believe the odor to be a distinctive one. Jungendfein (1937) tentatively ascribes it to poor treatment or to a cloth filtration. The latter, at least, is not applicable to our experimental wines, since they have not been filtered. Tentatively, we suggest that the odor is associated with a specific bacterial spoilage, probably lactic, in wines with fairly high pH, and this is being investigated.

Charcoal. This is very nearly the same odor as that of an overoxidized young wine. There is a marked increase in aldehyde. The effect, however, is not temporary as it is with oxidation alone. Traces of an odor of the charcoal remain, particularly if the charcoal is not of high quality. Some pink imported wines have a particularly noticeable charcoal odor, the charcoal having obviously been employed to secure the desired shade of color.

Bitter almond. Wines that have been over-blue-fined[4] may have an odor reminiscent of cyanide or bitter almonds. This is more often noted in high-acid, low-pH wines. Since the development of the sensitive Hubach test, few examples of this are encountered.

Mixtures of odors. In many cases more than one off-odor will be present. If they are present in appreciable quantities, they can be distinguished from each other, but if only traces of each are present it is often very difficult to identify them.

Odors—Desirable

All wines have a vinous odor. This is particularly strong in young wines and has been ascribed by Chauvet (1950) to an ester of lauric acid at a concentration of approximately 1 in 40,000. Wines that have no other odor are best described as *vinous.*[5] Most wine-grape varieties, however, contribute more than just a vinous odor to the wine. This is often sufficiently marked to be noticeable without being strong enough to be identified with a particular variety. The odor of such wines is properly described as *distinct.*

However, the finest wines are produced from varieties that give wines with a recognizable aroma. The unfermented grape itself may or may not have the characteristic odor, or it may be present in such small quantities as to be missed. When fermented in contact with the skins, however, the wine acquires this aroma, or the aroma becomes more easily distinguished. Varie-

[4] Blue-fining is the addition of potassium ferrocyanide to remove excess iron and copper from wine.

[5] Some experts maintain that every variety of wine grape produces a characteristic wine and that it is only our inexperience which prevents us from identifying it. Probably where a single variety is grown and is vinified under uniform procedures, the producers are able to identify it. The Chasselas doré wines of Switzerland are an example. This variety produces wines of almost neutral odor, yet Swiss connoisseurs can distinguish between the wines of different vineyards and vintages.

ties whose fruit, *if ripe*, can be identified by their odor include: Muscat Hamburg, Muscat of Alexandria, Orange Muscat, Muscat Canelli, Muscat Pantellana, Gewürztraminer, Pfeffer, Concord and many eastern hybrids, and sometimes Cabernet Sauvignon and Zinfandel. Varieties that seldom have such a distinctive aroma in the fruit but whose wine can be easily identified (if the fruit is mature and the wine properly vinified) include: Chardonnay, Semillon, Sauvignon blanc, White Riesling, Folle blanche, Sylvaner, Gamay Beaujolais,[6] Pinot noir, and Petite Sirah.

Pasteur classified the odor of a wine into primary aromas, which exist in the grape itself; secondary odors, derived from fermentation; and tertiary odors, which develop during aging. Mathieu (1902) distinguished between odors derived from the grape directly (muscat, etc.) and those formed during fermentation from non-odorous substances in the grapes. He also noted the odors developed during aging in the wood and aging in the cask. Bouquet development in the bottle, however, depends on the character of the wine before bottling. In this publication, the term "aroma" refers only to an odor derived from the grape. The secondary odors are described as fermentation odors or as bouquet. There is reason to believe that primary odors are derived directly from the fruit, or at most are only slightly modified during fermentation. Alcohol may also have some influence on the detectability of odors. However, Hennig and Villforth (1942) state specifically that their pentane extracts clearly revealed the aroma of the variety from which they were extracted (Sylvaner, Traminer, Müller-Thurgau) and that they contained little alcohol.

As to the chemical nature of these odors we have little information. The foxy aroma of the Concord variety (*Vitis labrusca* Bailey) was attributed to methyl anthranilate by Power (1921) and Power and Chesnut (1921). They found none of the ester in six varieties of *V. vinifera* nor in several of *V. rotundifolia*. Catawba, a *V. labrusca* × *V. vinifera* hybrid, was also free of the ester, but it was present in other similar hybrids. However, Catawba wines vinified here (Davis) are clearly "foxy." The presence of methyl anthranilate in Concord grape juice has been confirmed recently by Holley *et al.* (1955). They reported 90 per cent of the volatile organic material to be ethyl alcohol. Other constituents were methyl alcohol, methyl and ethyl acetate, acetone, acetaldehyde, acetic acid, and methyl anthranilate. The essence employed represented 1.2 per cent of the original volume and contained 0.033 mg per ml. The original Concord grape juice, therefore, had about 0.4 mg per liter of methyl anthranilate. The high percentage of ethyl alcohol in the volatile organic material simply means that dichromate or other oxidation procedures for evaluating the amount of odor do little more than measure the amount of ethyl alcohol.

Robinson *et al.* (1949) reported methyl anthranilate in Concord grapes to increase from less than 1 mg per liter in early September to over 5 mg per liter in mid-October. Sale and Wilson (1926) reported 0.0 to 3.8 mg per liter of this ester in various "Eastern" musts. Scott (1923) found 0.80 to 1.49 mg per liter in commercial Concord grape juice but only 0.08 to 0.40 in com-

[6] But not the so-called Napa Valley Gamay. The Gamay from the Beaujolais region in France is reported by Chauvet (1950) to contain *p*-hydroxyphenylethyl alcohol (probably derived from phenylalanine), which has a roselike odor.

mercial Catawba grape juice. It is possible the Catawba juices contained some Concord juice, since Power and Chesnut (1921) did not find methyl anthranilate in Catawba. However, Catawba must have a similar type of compound, as its juice is very "foxy."

In *Vitis rotundifolia* grapes Kepner and Webb (1956) found the volatile essence to consist of ethyl, *n*-butyl, *n*-hexyl, and β-phenylethyl alcohols, and acetate, laurate, and isopropyl esters. Probably present were methyl alcohol, *n*-hexanal, 2-hexenal, and acetal. When a fermentation technique was employed, more compounds were found but some were produced by fermentation. The revealing compound is the β-phenylethyl alcohol, which has a rose-like odor reminiscent of that of *V. rotundifolia* grapes. This confirms the finding of Power and Chesnut (1923) that methyl anthranilate is not present in this species.

Hennig and Villforth (1942) made a complete analysis of a wine which contained a good deal of Müller-Thurgau (*V. vinifera*) grapes. This variety has a moderately distinctive aroma in the wine (it is a cross of Sylvaner and White Riesling). They found with certainty or at least a strong possibility: formaldehyde, acetaldehyde, propionaldehyde, vanillin, cinnamaldehyde, acetone, acetylmethylcarbinol, acetal; the following alcohols—methyl, ethyl, isopropyl, isobutyl, isoamyl and α-terpineol, *n*-propyl, *n*-heptyl, *sec*-nonyl (2-nonanol); and the following acids—formic, acetic, propionic, *n*-butyric, caproic, caprylic, capric, and lauric. Their tests suggest that methyl ketone, caproaldehyde (and higher aldehydes), benzaldehyde, furfural, and pelargonic and myristic acids are also present. Hennig and Villforth suggest that the alcohols and acids are present mainly as esters, since little free acid or alcohol was present.

Haagen-Smit *et al.* (1949) studied the Zinfandel, likewise a variety of *V. vinifera.* They reported the following compounds (mg per kg of fruit): ethyl alcohol, 244; acetaldehyde, 1.80; acetic acid, 0.0053; *n*-butyric acid, 0.003; *n*-caproic acid, 0.0015; glyoxylic acid, 0.118; *n*-butylphthalate, 2.250; leaf aldehyde, 0.327; sulfur, 0.004; acetylmethylcarbinol, 0.013; waxy substances, 0.024; and carbonyl compound, 0.025. This does not represent all the odorous material of the Zinfandel nor does it explain the unique odor of its wine; moreover, the source of the *n*-butylphthalate has been questioned. Recently Cordonnier (1956) has reported the muscat odor to be partially due to linalol or a derivative of it.

Webb and Kepner (1957) have investigated the volatile constituents of another *V. vinifera* variety, the Muscat of Alexandria. They found (in mg per kg of grape berries): methanol, 3.7; ethanol, 111.0; *n*-butanol, 0.03; 3-methylbutanol, 0.01; *n*-hexanol, 0.49; cis-3-hexenol, 0.26; acetaldehyde, 0.85; *n*-hexanal, 0.03; 2-hexenal, 0.05; 2-butanone, 0.01; 2-pentanone, 0.01; methyl acetate, 0.08; ethyl caproate, 0.04; and a mixture of butyrate, valerate, caproate, caprylate, caprate, laurate esters, and ethyl esters, 0.16. Acetals were also present. They suggested that the failure to find geraniol, terpineol, limonene, or linaloöl, as reported by Cordonnier (1956) in muscat grapes, may be due to their decomposition in the long heating in the stills used for fractional distillation in their study.

Wines made from grapes attacked by the mold *Botrytis cinerea* have a strong and distinctive odor the source of which has not been identified. In

wines aged in oak casks or fermented with the grape seeds the presence of vanillin has been suspected for many years. Göpfert (1953), however, did not find it in five European dessert wines. On the other hand, Grohmann and Muhlberger (1954), using an improved technique, found from 0.02 to 0.25 mg of vanillin per liter in 22 European dessert wines, the higher amounts being found in wines that had been aged in the cask the longest, suggesting perhaps that some of the vanillin is extracted from the wood.

Fermentation might develop odors and tastes through various mechanisms: deamination and decarboxylation of amino acids to yield higher alcohols, hydrogenation of other hydrogen acceptors than acetaldehyde, acyloin syntheses of carbinols ($RCHO + CH_3CHO \rightarrow RCHOH\ COCH_3$), Cannizzaro reactions involving acetaldehyde and fermentation intermediates to produce higher acids ($2RCH_2CHO + H_2O \rightarrow RCOOH + RCH_2OH$), and formation of acids as the yeast utilizes ammonium ion or as a result of sugar assimilation (such as produces succinic). The acids produced may later react to form esters. Direct ester formation is pronounced with certain yeasts. Release of hydrogen sulfide and mercaptans through the reduction of free sulfur or sulfur-containing amino acids should also be noted. Of course, alcohols and aldehydes have profound influences on the odor of all types of wine. The question of how much of the odor of a table wine is due to the yeasts has been reviewed by Amerine and Joslyn (1951), Castor (1954), Crowther (1951–1952), Schanderl (1950), Ventre (1935), and others. As Crowther points out, the results are not easy to interpret. This in itself is some evidence that the influence of the yeast is not clear-cut or profound. Certainly among the various strains of *Saccharomyces cerevisiae* var. *ellipsoideus* it is difficult to find marked differences when the various strains are employed to ferment the same must. Ventre (1935) obtained yeasts from fruit of the Beaujolais, Burgundy, Champagne and Médoc regions of France, which he inoculated into an artificial medium and also into Aramon musts sterilized by heat and by sulfur dioxide. The odor of the wines was suggestive of the region of origin of the yeasts, but the chemical analyses did not reveal any characteristic differences. In some Italian fermentations mixed cultures are common. The yeast flora, particularly where *Kloeckera* sp. and *Torulaspora* sp. are present, may contribute to the character of the wine, but no systematic field or even pilot-plant studies have been made with adequate organoleptic and chemical controls.

Castor (1954) reported *Schizosaccharomyces pombe* to produce an unusual and not unpleasant aromatic odor. He noted that the flavor profiles for the bakers' (*S. cerevisiae*), Burgundy, and Tokay (*S. cerevisiae,* var. *ellipsoideus*) strains, and for *Schizosaccharomyces, Saccharomycodes,* and *Brettanomyces* were similar.

The situation with respect to sherry is entirely different. Here the film stage of the *flor* yeast is responsible for the chief odor characteristic of the wine. Acetaldehyde, acetal, and probably diacetyl are among the immediate products. But the odor of older *flor*-yeast wines is very complex. Products of the autolysis of the yeast deposits may be partially responsible.

Methyl alcohol. Usually less than 1 per cent of the total alcohol content of a wine is methyl. It is very doubtful that it has any significance for the odor of wines. Its esters, however, have agreeable odors and may be im-

portant. Amerine (1954) summarized the analytical results to date and showed averages of 0.016 to 0.181 gm of methyl alcohol per liter in various types of wines.

Ethyl alcohol. This is one of the few common chemicals that has odor, taste, and feel. The alcohol content of wines varies from as low as 8 per cent (by volume) to 21 per cent or more in vintage ports. The thresholds for these have been reported by Parker (1922) as: 0.44 per cent (by weight) for odor, 14.0 per cent for taste, and 25.0 per cent for the common chemical sense.

The effect of alcohol in moderating the influence of other substances in the wine has been emphasized by Ribéreau-Gayon (1950). He notes that if the alcohol is separated from the wine by vacuum distillation, the taste of the dealcoholized wine is very acid. This is probably one reason why some low-alcohol white and red wines are so acidulous. For an odor threshold Berg *et al.* (1955*b*) reported 0.004–0.0052 gm of alcohol per 100 ml. It is not clear whether or not taste was eliminated in their tests, even though the tasters were told to judge on odor. At various concentrations of alcohol the following differences for detection were necessary: at 4.0 gm per 100 ml, 1.2 to 1.5 gm per 100 ml; at 8 gm per 100 ml, 1.6 to 2.4; at 12 gm per 100 ml, 2.4 to 3.2; and at 16.0 gm per 100 ml, 2.8. Hinreiner *et al.* (1955*b*) also studied the effect of sucrose on the threshold for ethyl alcohol and on the minimum detectable concentration. Their results were as follows:

Detectable Concentration Difference in Per cent Alcohol

Ethyl Alcohol Per cent	Sucrose			
	0%	5%	10%	15%
0	0.005	0.15	0.35	0.45
10	2.0	3.0	4.0	4.0
15	3.0	4.0	4.2	4.2

Thus sugar increases the threshold for alcohol and also the minimum-detectable-difference concentration, and these effects increase with increasing sugar and alcohol concentrations. In a similar manner they demonstrated that the threshold for alcohol increased with increasing sucrose in an acid solution, but the acidity had some effect on diminishing the effect of the sugar. With wines, however, Hinreiner *et al.* (1955*a*) reported a detectable difference concentration of 4 per cent by volume for a dry white table wine of 11 per cent alcohol and a red table wine of 12 per cent alcohol. Sweetening the red wine to 10 gm per 100 ml of sucrose did not increase the difference threshold.

Higher alcohols. The amount of higher alcohols (fusel oil) present in wines varies markedly. In a few cases they can be recognized as such in the wine. This is particularly true of dessert wines that have been fortified with grape spirits high in fusel oil. Guymon and Heitz (1952) reported the following amounts in California wines (gm per liter):

Type	Minimum	Maximum	Average
White table	0.162	0.366	0.250
Red table	0.140	0.417	0.287
Dessert	0.156	0.900	0.374

The higher alcohols, like the other alcohols, may be important not only because of their own odor but also because of their solubility and volatility in relation to other odorous materials (as an off-odor, see p. 498).

Aldehydes. Acetaldehyde is the main aldehyde present in wines. Amerine (1954) reports the following amounts in California wines (mg per liter):

Type	Number of samples	Minimum	Maximum	Average
White table.........	480	9	92	54
Red table...........	170	5	494	46
Dessert.............	142	15	217	87

Spanish sherries (film yeast) had 90 to 500 mg per liter, average 218. Berg *et al.* (1955*a*) found that their panel could distinguish 0.00013–0.00015 gm per 100 ml of acetaldehyde in water. Their experienced tasters were much more sensitive to acetaldehyde than their inexperienced group. Hinreiner *et al.* (1955*a*) reported a difference threshold of 125 ppm in a dry white wine and 100 ppm in a dry red table wine. The original acetaldehyde content of the wines was not given. Aldehydes may also appear during aging (see p. 496). Higher aldehydes and acetals also may more appropriately be discussed under aging. Zellner (1927) suggested that even in non-*flor*-sherry wines, aldehydes and ketones play a part in the bouquet, the kind being more important than the quantity. He also suggested that a quantitative measurement of the carbonyl group would be of value in characterizing wines of a particular vintage, etc. (for aldehydes as an off-odor, see p. 496).

Acetal. This seems to be a component of film-yeast sherries, but owing to analytical difficulties there are few data. Amerine (1954) summarized the data as showing only traces. Acetal is in equilibrium with acetaldehyde at low pH's, but only a small percentage of the aldehyde is converted to acetal in wines. Webb and Kepner (1957) reported acetals in fresh muscat grapes.

Hydroxymethylfurfural. The presence of this compound in baked wines has been demonstrated by various investigators. It is produced by dehydration of levulose. Amerine (1948*b*) found 32 to 305 mg per liter in six California baked sherries. Wines made from raisins and wines to which reduced must has been added may contain even larger amounts. Grape concentrate is also reported to have an appreciable amount of hydroxymethylfurfural, and so, presumably, wines made with concentrate would also. While this substance has a distinct odor, it is probably only a part of the caramel-odor complex.

Esters. Both neutral and acid esters occur in wines. All of the acid esters are nonvolatile, but the neutral esters may be volatile (ethyl acetate, ethyl lactate, etc.) or nonvolatile (ethyl tartrate, ethyl malate, etc.). In California wines the average total neutral and volatile neutral esters were summarized by Amerine (1954) as follows (mg per liter as ethyl acetate):

Type	Number of samples	Neutral esters Total	Neutral esters Volatile
White table.............	79	209	73
Red table...............	60	400	112
Dessert.................	124	333	104

The role of esters in the aroma and bouquet of a wine has been the subject of much controversy on the basis of too little data. The classic enologists—

Pasteur, Berthelot, Kayser, and Ventre—assigned an important role to the esters. Hennig and Villforth's (1942) work with Müller-Thurgau variety also seemed to indicate that the esters were important. Peynaud (1937), however, found evidence that ethyl acetate was the only ester important in the odor. This appears to be too great a simplification, although for many wines it is essentially correct. Further evidence of its correctness is the rather mild odor of pure solutions of many of the possible esters of wines: ethyl tartrate, malate, and succinate have very mild odors. The role of esters of lauric acid has been mentioned (p. 499).

For ethyl acetate Berg *et al.* (1955*a*) reported a threshold in water of 8–11 mg per liter. With an initial concentration of 30 mg per liter, a difference of 40 was necessary for detection. Their experienced group of tasters was more sensitive to this odor than their inexperienced. Hinreiner *et al.* (1955*a*) found a difference threshold of 175 mg per liter in a white table wine and of 200 mg per liter in a red table wine. The original total volatile neutral ester content of the wines was reported as 212 and 308 mg per liter respectively. These very large difference thresholds need to be studied further. Diacetyl is present in even smaller amounts, 0 to 6 mg per liter, but because of its very distinctive odor it may be a factor in the odor of some wines, particularly very old dessert wines.

Tastes—Undesirable

Most discussions on tastes list a large group of so-called undesirable tastes. We have already defined many of these under odors and have placed them in the category of undesirable odors, even though some are perceived only in the process of tasting or taking into the mouth. However, a few true tastes are definitely undesirable.

Sulfuric. Wines of high sulfate content have a bitter taste. If the sulfuric acid content is high enough to reduce the pH below 3, the wine may have a dry, hot, excessively acid taste.

Bitter. Red wines containing more than 0.25 gm tannin per 100 ml tannin will usually taste bitter. A number of compounds other than tannins have bitter tastes. Among these are sulfates, unknown constituents of green grapes, stems, over-heated pomace, etc. At times we have tried to differentiate a bitter taste from a rough sensation. This may be possible, but so far we have no explanation for the differences in bitter sensation for wines of equal tannin content. However, terms such as *slightly bitter*, *bitter*, or *very bitter* are used interchangeably with *slightly rough*, *rough*, or *very rough*.

Sweetish. Grazzi-Soncini (1892) defined this as an undesirable sweet taste in table wines. The term appears useful, especially for table wines of low acidity where the sweet taste gives a rather sickening impression.

Flat. This refers to wines with insufficient acid taste. It is applied particularly to young table wines, which lack acidity. In extreme cases it is apparently related to the rubbery odor (p. 498) associated with a high pH.

Acidulous. This taste is associated with excessive total acidity. It is sometimes described as a steely, hard, green, or even metallic taste. Wines of low alcohol and extract contents that are also high in total acidity and low (below 3) in pH are particularly disagreeable and should be marked as *acidulous* (for the green odor see p. 497).

Tastes—Desirable

Acid. The primary acids responsible for the acidity of wines are tartaric, malic, lactic, acetic, and succinic. Of these, acetic has a recognizable odor, while lactic in the concentrations found in wine has only a slight odor. The other acids are odorless. In the descending order of their effect on the pH of the wine the acids are tartaric, malic, succinic, lactic, and acetic. The relative proportions of the acids present in the finished wines depend on the amounts and ratio of tartrates and malates in the must, the amount of bacterial action during and after the fermentation, and the temperature of storage or the amount of tartrates removed or changes produced by ion-exchange treatment. The acid taste is due in part to the hydrogen ion content of the liquid (Paul, 1917). One would expect this to be reflected in a direct relationship between pH and the degree of acid taste, but this is not the case. Rather, the acid taste appears to be a function of the pH and the titratable acidity. The buffer capacity of the wine, its sweetness, and other components of the extract also influence the acid taste.

Berg *et al.* (1955*a*) found the following thresholds (in gm per 100 ml of water): sulfurous acid, 0.0011; tartaric acid, 0.0024–0.0027; citric acid, 0.0023–0.0025; potassium acid tartrate, 0.0075–0.0090; lactic acid, 0.0038–0.0040; malic acid, 0.0026–0.0030; and succinic acid, 0.0034–0.0035. For difference detection they reported in aqueous solution 0.05 gm per liter at 0.25 per cent *d*-tartaric acid, 0.10 for 0.30 per cent potassium acid tartrate, 0.07 for 0.21 per cent citric acid, and 0.05 for 0.23 per cent *l*-malic acid.

In the studies of Hinreiner *et al.* (1955*b*) sucrose (to 10 per cent) did not influence the threshold for acid in aqueous solutions. Ethyl alcohol, however, did increase the acid threshold, and the effect was greater if sucrose was present. Tannin also seemed to increase the minimum detectable differences for acid in the absence of sugar. Sugar in the absence of tannin did not have this effect. The presence of sugar tends to minimize the effect of tannin on the minimum-detectable-difference concentration of acid. In a white table wine, Hinreiner *et al.* (1955*a*) reported a difference threshold for tartaric acid of 0.15 gm per 100 ml. The original acidity was 0.74 gm per 100 ml as tartaric and the pH 3.3. When the wine was sweetened with 1 gm per 100 ml of sucrose, the difference threshold was not increased.

Ventre (1931) studied the acid taste of four red and four white French wines. The tasters (number not stated) were asked to rank the wines in the order of their acidity. The results were as follows:

Taste	Alcohol, per cent	Total acidity, per cent tartaric	Tartrates, per cent potassium bitartrate
	RED WINES		
Least acid	10.6	0.81	0.29
Some acid	10.1	0.89	0.31
More acid	8.3	0.80	0.31
Most acid	9.0	0.75	0.43
	WHITE WINES		
Least acid	10.2	0.89	0.20
Some acid	11.0	0.86	0.21
More acid	9.5	0.84	0.41
Most acid	8.9	0.77	0.38

He interpreted these data as indicating that the tartrate ion was responsible for the acid taste. However, the pH and malic and lactic acid contents were not given nor was the influence of the alcohol content considered. Ventre did believe that pH and the acid taste were closely related. There is no conflict between these two ideas, since tartaric acid is the strongest of the organic acids present. In another case, however, he found malic acid highest in the more acid wine.

We prefer to use the terms *low, medium,* and *high* acid for degrees of acid taste. While *fruity* and *tart* have implications of acid taste to some, there is much confusion as to their exact meaning. *Fruity* seems to indicate a pleasant acid taste, but it may involve odor as well as taste. *Flat,* meaning a lack of acidity, is a useful word. *Green* is a special term aptly applied to wines of unripe grapes. Usually the total acidity exceeds 0.85 per cent in such wines. Sometimes a special *green* aroma may be noted (see p. 497).

Tannin. There is little bitter taste in white wines, since they contain less than 0.05 per cent tannin. The astringency of red wines is due to tannins and is an important factor in judging their quality. Amerine (1954) reported 0.10 to 0.38 per cent (average 0.21) in 282 California red table wines. The tannin content of red wines gradually decreases during aging.

Berg *et al.* (1955*a*) reported a threshold for grape-seed tannin as 0.020 gm per 100 ml of water. Their experienced tasters were somewhat more sensitive to tannin than their inexperienced group. Hinreiner *et al.* (1955*a*) found a difference threshold of 0.10 gm per 100 ml for a white table wine and 0.15 gm per 100 ml for a red table wine. The tannin content of the wines was 0.02 and 0.20 gm per 100 ml respectively. Increasing the sugar content to 5 and 10 gm per 100 ml by adding sucrose did not change the difference threshold of the red wine. These high difference thresholds are higher than those found in red wines. The higher difference threshold here may be due to differences in the tannin that was added and that naturally present in grapes. The distracting influence of the other constituents may also be a factor.

In many wines an astringent taste is not a negative quality factor. Some tasters, however, are more sensitive to bitterness than others. A certain astringency seems particularly desirable with red wines when they accompany certain foods. It has also been suggested that bitter-tasting drinks may increase the appetite (or postpone a feeling of satiety). It is noteworthy that experienced tasters are more likely to prefer wines with a higher astringency. There is also a possiblity that wines which are originally high in tannin may age more satisfactorily for longer periods of time, with release of desirable odorous constituents, than low-tannin wines. The anti-oxidative property of tannins should also be considered.

Sugar. Sweetness in wine is not due solely to dextrose and levulose but also to glycerol, 2, 3-butylene glycol, etc. Dextrose is much less sweet than levulose. When sucrose is added to finished wines, it is usually rapidly inverted and hence has the sweetness of invert sugar.

In mature grapes, the dextrose/levulose ratio approaches one according to Amerine and Joslyn (1951), but it may be greater than one for overripe grapes and less than one for green grapes. Moreover, different varieties of yeasts ferment dextrose and levulose at different rates, and the percentage

of each present thus depends on this initial concentration, the total amount of sugars, and the extent of fermentation. For a discussion of the possible importance of the dextrose/levulose ratio see Amerine and Thoukis (1958).

Berg *et al.* (1955*a*) reported thresholds in water of 0.13–0.15 gm per 100 ml of levulose, 0.40–0.44 for dextrose, and 0.38–0.44 for glycerin. With various concentrations of dextrose in water, the following differences in concentration were necessary for detection: at 1 per cent, 0.4–0.7; at 5 per cent, 0.6–0.8; at 10 per cent, 0.8; and at 15 per cent, 1.0–1.1. The effect of sugar in reducing the sharpness of the acid taste is well known. With low-alcohol high-acid wines the presence of sugar can be of critical importance to the palatability of the wine. The sugar of many German wines has obviously been retained with this in mind.

The influence of alcohol, acids, and tannins on the sweetness of sucrose (sucrose having the same threshold as a mixture of 50 per cent dextrose and 50 per cent levulose in these tests) was also studied by Berg *et al.* (1955*b*). Ethyl alcohol enhances the sweetness of sugar solutions from 1 to 15 per cent sucrose as indicated both by the threshold and the minimum detectable difference. In non-alcoholic solutions the effect of pH in the range 2.55 to 3.40 was negligible both on the sugar threshold and on the minimum detectable difference. However, these values were higher than those of sugar-water solutions. Tannin also increased the threshold level for detection of sugar and the minimum detectable difference. In wines Hinreiner *et al.* (1955*a*) reported the following difference thresholds at various concentrations of sucrose:

Sucrose level	Minimum detectable sucrose increment in	
	White wines	Red wines
gm/100 ml	*gm/100ml*	*gm/100ml*
0	0.9	0.8
1	0.8	0.7
5	1.1	1.0
10	1.7	1.6

These differences are greater than for aqueous solutions but similar to those for water-alcohol-acid solutions. Experienced tasters seem to be able to detect somewhat smaller differences than these. Perhaps the high values may be due to the use of a triangular rather than some type of paired test (see p. 529).

Glycerin. About 0.3 to 1.5 per cent of wine is glycerin, which has a sweet taste. Berg *et al.* (1955*a*) reported a threshold of 0.38–0.44 per cent in water. Hinreiner *et al.* (1955*b*) found that acids and alcohol increased the glycerin threshold. An acid solution of about pH 3.4 increased the threshold to about 1.5 gm per 100 ml and 10 per cent alcohol raised it to 1.0 gm per 100 ml. In wines, Hinreiner *et al.* (1955*a*) reported a difference threshold of 0.9 gm per 100 ml in a dry white table wine and 1.3 gm per 100 ml in a dry red table wine. The original glycerin content of the wines was 0.3 and 0.8 gm per 100 ml respectively.

The Common Chemical Sense

Much of the objectionableness of sulfur dioxide (p. 494) is due to its effect on the common chemical sense rather than the olfactory nerve. The sneezing associated with consumption of wines high in free sulfur dioxide is almost certainly associated with the latter effect.

Alcohol also has an influence on the common chemical sense, and it is difficult to dissociate this effect from its effect on the olfactory nerve and on taste. However, wines of low extract and high alcohol probably "burn" because of their high ethyl alcohol content.

Burnt. In addition to the caramel odor of baked wines, there is often present in overbaked wines a pungent principle which affects the common chemical sense. This is reflected in the tendency to sneeze.

Hot. This should be carefully distinguished from the preceding odor. There are apparently two types. First, table wines of excessively high alcohol content produce a hot or burning sensation. Second, dessert wines with a high amount of aldehyde and related compounds are said to be *hot*. In both cases taste, as well as the common chemical sense, seems to be involved.

Kinesthetic and Tactile Sensations

Kinesthetic and cutaneous sensory impressions have often been confused with taste, probably because both are sensed in the mouth. The important kinesthetic impression is due to the extract content, while carbon dioxide is associated mainly with the tactile sensation.

Wines of high extract content have a feel that is quite different from those of low extract. Alcohol and sugar can also influence this sense, but it is normally true that if the alcohol content of a table wine is high, the extract will also be high. The following terms have been used. Wines of low extract and alcohol, often accentuated by high total acidity, are said to be *thin*. *Unbalanced* or *unharmonic* are related terms; the former, however, seems to apply more to excessive acidity than to lack of body. Wines of high body are often said to be *full*, *round*, or *rich*, implying a mouth-filling sensation. Such wines are *soft* or *smooth*, a sensation that may be due to the sweet taste of glycerin or sugars, or to the aging and decrease in bitterness due to oxidation and precipitation of tannin substances.

It is possible that the very *dry* taste of old table wines may be a tactile rather than a taste sensation.

The role of carbon dioxide in the taste and odor of wines has not been studied adequately. This is curious, since sparkling wines are an important part of our wine production. The fact that carbon dioxide accentuates the bitter taste is recognized by most authorities. Red sparkling wines, such as sparkling burgundy, are sweetened much more than their white counterparts, ostensibly to reduce the bitter taste of the tannin plus the carbon dioxide. Carbon dioxide has a variety of effects on the organoleptic examination of wines. Certainly it has an acid taste. Its role in preventing oxidation has not been fully studied. Contrary to what we should imagine, sparkling wines, even with considerable carbon dioxide, are not immune to oxidation, or what we normally call oxidation. Some old sparkling wines of low pressure, but still with some pressure, turn brown and have an oxidized odor.

Probably this happens much more slowly than if there has been no pressure, but the reaction is curious. Since most of these wines are sweetened, some form of the Maillard reaction may be partially responsible for the brown color. Of course, diffusion of oxygen, even against a pressure gradient, is not impossible, but it is certainly very slight. In the case of wines produced by fermentation in tanks, considerable oxygen pick-up may occur during transfer, filtration, or bottling unless a counterpressure of nitrogen gas is employed. (Unfortunately, the more logical use of carbon dioxide for the counterpressure is not permitted in this country.)

Carbon dioxide, however, seems to have its primary effect on the tactile sensation. The *gassy* sensation of many sparkling wines makes a tactile impression often described as prickly or biting.

In Europe the effect of carbon dioxide in increasing the desirable acid taste, or in increasing the odor, in widely recognized. Engel and Pernot (1937) reported as much as 200 cc of carbon dioxide per liter of mature white and red table wines. Young wines had up to 300 cc. In Switzerland, white table wines are often bottled with added carbon dioxide. *Pétillant,* or gassy, wines are found in many local areas. Those wines lose much of their attraction when they lose their gassiness—as they often do when transported through or into warm regions or to cellars of variable temperature. This is probably why some European wines are reported not to travel well. Furthermore, the release of carbon dioxide brings both the desirable and undesirable odorous constituents of the wine to the nose much more quickly.

General Impression

Although most of the organoleptic properties of a wine can be described in terms of odor and taste, of their effect on the common chemical sense, and of the kinesthetic or cutaneous reactions they produce, the taster frequently finds it necessary to integrate his impressions. This over-all impression is not unimportant even though it obviously represents a synthesis of the senses just mentioned. Some of the remarks made about wines by experienced tasters can therefore only be classified under general impressions. The elusive *sève* of the French apparently belongs in this category. It expresses the combined sensation of taste and odor. Got (1953) defines it as a measure of finesse, silkiness (*soyeux*), and distinction of the wine. In beer-tasting Kutter (1948) specifically separates the after-impression from the first impression. As will be seen later, in some of the score cards this over-all impression is identified as a special characteristic. Baker's (1954) statistical study also seems to suggest that there is a general quality factor.

Among the terms used are *harmonious* or *unharmonious*, and *young, mature,* and *aged*. *Harmonious* describes a pleasant balance of all the components of a wine, which gives the taster a particularly pleasant sensation. This term should be employed discreetly, however. *Unharmonious* describes a lack of balance of the constituents of a wine, which results in an unpleasant over-all impression. This term likewise should be used with care, since the cause of the lack of balance should be noted under the individual organoleptic characteristics. *Young, mature, aged* obviously have to do with changes taking place during aging. *Young* wines are characteristically fruity and even gassy. For certain table wines this is not undesirable. A *mature* wine

is one that has lost its fruity-yeasty-gassy character and has a full vinous tone. It is properly applied to table wines that are ready for bottling. *Aged* signifies further development. Table wines that have been bottled for six months to five years or more acquire a bottle bouquet which is quite characteristic. Dessert wines likewise age in the bottle, but they are also said to be aged when they have acquired some bouquet from aging in the wood.

IV. DESCRIPTION OF WINE TYPES

Thousands of different types of wines are made in different parts of the world. These differ from each other in color, taste, and odor. To the uninitiated the differences are confusing. Even the expert cannot hope to master the minute differences within types for all of these wines. It is not our purpose here to attempt a complete classification of the wine types presently produced in California. We seriously doubt if such a classification can be attained at the present time in this state. For one thing, the commercial usage of nomenclature is hopelessly confused among members of the industry. Furthermore, the California wineries are obviously in a state of change with respect to their concept of what constitutes certain types.[7] The classification proposed is strictly for wines produced in California, and type names such as chablis, port, sherry, as defined here, are obviously not valid equivalents for similarly named European types, nor, in some cases, are they completely equivalent even throughout the California wine industry. European climatic effects, varietal composition, and methods of wine making differ so widely from those prevalent in this state that no comparisons seem relevant.

The definitions of types given by the Wine Institute (1956) for the guidance of judges of fairs are not very specific and obviously represent a compromise between various points of view. The classification of Berg and Webb (1955) is certainly better than that of the Wine Institute but has not been accepted by the industry.

The classification given here has been condensed to provide for distinctions which we believe can be most generally recognized. Those who need further subdivisions can easily provide them as needed. Some producers, for example, maintain clear and uniform distinctions by blending between types, such as claret and burgundy or chablis and rhine. Unfortunately, others have different standards or none at all. Furthermore, it is questionable, with our present large use of table grape varieties and of grapes of low varietal distinction, whether such distinctions are universally possible. A reduction in the number of generic-type names would be highly desirable.

Table Wines

Table wines contain 10.5 to 14 per cent alcohol (10 to 14 per cent for whites). There are two possible major groupings: by color or by sugar content. Since the color groupings are more easily defined, we will classify all table wines as white, *rosé* (pink), and red. The color specifications given on pp.

[7] California producers are not alone in this. The German wine trade has changed the German wines from their traditional dryness to more than a touch of sweetness since World War II. The Bordeaux concept of claret is likewise changing from a slow-maturing to a rapidly aging red wine.

491, 492, and 493 may be used to obtain the limits, but few people will have any difficulty in distinguishing white and *rosé*. Some wines which should be red may be just pink. At present there are no standards to prevent the low-color wine from being sold as a red rather than a *rosé*. The opposite case seldom, if ever, occurs. Within each group, sugar content and varietal odor will be used as the primary points of differentiation.

White. Most California white table wines do not have a recognizable varietal character. *California dry white table wine* seems the most valid appellation for the nonvarietal types. These are wines with light to medium yellow color and no distinguishable sugar content. This group includes wine types labeled California chablis, rhine, dry sauterne, and white chianti. Types with varietal names such as Ugni blanc, Golden Chasselas, or Veltliner, rarely have sufficient varietal character to be distinguishable. All other white nonvarietal table wines may be designated as *California sweet white table wine*. These contain 0.5 to about 7 per cent sugar. The sulfur dioxide content should not be excessive. This group includes types such as sauterne, chateau-type sauterne, and mellow white table wines, usually with a proprietary label.

A great many varietal types are produced in California—mostly in small quantities, but nevertheless important as far as the reputation of the industry is concerned. The amount of varietal character in any given wine depends on the source and maturity of the grapes and on the fermentation, blending, aging, and finishing procedures employed. Wines that do not have a recognizable character even though produced from only one variety should not, in our opinion, be given a varietal label. In approximate descending order of the distinctiveness of their varietal character they are: Gewürztraminer, dry Sauvignon blanc, dry Semillon, Chardonnay, White Riesling, Sylvaner, French Colombard, Folle Blanche, Chenin blanc (erroneously termed White Pinot) and Grey Riesling. Sauvignon blanc and Semillon are produced without the prefix "dry" (about 1 to 4 per cent sugar) and also as sweet Sauvignon blanc and sweet Semillon (about 4 to 7 per cent sugar).

Rosé. The type name *California rosé* includes both dry and sweet (0.5 to about 5 per cent sugar) wines. This confusion could be avoided if the dry *rosés* were so labeled. Varietal *rosé* types, such as Grenache, Gamay, Pinot noir, and Cabernet, are also produced. The Grenache and Gamay are often barely, if at all, distinguishable in aroma from wines labeled simply *rosé*. At best, the Grenache has more orange tint and less acid than the Gamay. The varietal character should be easily recognized in the Pinot noir and Cabernet *rosés*.

Red. Most of the California red table wines are blends and hence have no distinguishable varietal aroma. They can best be designated *California dry red table wine*. This includes such types as California claret, burgundy, and chianti. In addition, wines labeled after the varieties Carignane (sometimes labeled Mourastel) and Petite Sirah (also called Durif) are seldom strong enough in varietal aroma to be recognizable. California sweet red table wines (1.5 to 3 per cent sugar) usually bear proprietary labels.

The varietal types, in descending order (approximate) of the distinctiveness of their varietal character, include Cabernet, Zinfandel, Pinot noir, Barbera, Gamay, and Red Pinot (for Pinot St. George). All should be dry.

Dessert Wines

Dessert wines in California contain 19.5 to 21 per cent alcohol. They are distinguished on the basis of color, varietal odor, or odor derived from the process of production.

White. The main nonvarietal and nonprocessed white dessert wine is Angelica. This is a wine with a bland, high-sugar taste. It may, however, be aged to acquire a distinctive bouquet.

The only white varietal types commonly produced are muscatel (usually made of Muscat of Alexandria) and Muscat Canelli (more commonly labeled Muscat Frontignan). The muscat flavor should be pronounced and free of oxidation or of excessive woodiness. Though many are amber, these wines at their best are yellow or gold.

The most common processing for dessert wines in this state is baking. The baked type is called sherry. Dry (or cocktail) sherry should have less than 1.5 per cent sugar, sherry 1.5 to 4 per cent, and sweet (or mellow or cream) sherry over 4 per cent sugar. The baked character should not be too pronounced, and a slight oakiness is desirable.

No uniform procedures for producing madeira or marsala prevail in the industry, and production is very small. Since California sherry is baked, madeira as a type name is unnecessary. Marsala, as a type containing reduced must, might have a place. It should be amber, have a distinct burnt odor, and contain 8 to 12 per cent sugar if it is to resemble its European counterpart.

White dessert wines aged under a film yeast are commonly called *flor-type sherry*. Dry-*flor*, *flor*, and sweet-*flor* types are produced. The recommended percentage of sugar is as for sherry. No baked wine should be blended into these wines if the distinctive flor character is to be maintained.

Amber-red. A blended type, California tokay, with an amber-red color is also produced. It is a blend of angelica, sherry, and port. The sherry character should not predominate. This type has no inherent connection with the table grape variety Flame Tokay.

Red. Several red dessert wines are now produced in this state. The varietal types include red muscatel and Tinta (for Tinta Madeira or Tinta Cão or blends). Nonvarietal types include California ruby and tawny port and wine labeled simply "port." Tawny ports should be red with a distinct amber tint; they should be free of baked odor. Ports and ruby ports should have a moderate red color and no amber tint.

Sparkling Wines

Wines containing excess carbon dioxide may be divided into two legal classes, sparkling and carbonated. Both of these carry a high charge of gas, 3-5 atmospheres of carbon dioxide (45–75 lbs. per sq. inch of pressure). There is, in addition, a third type now classified as a wine containing excess carbon dioxide but which may be reclassified as a still wine because of its extremely low charge (less than about 7 lbs. per square inch of pressure).

Sparkling wines are produced by a refermentation process, that is, the fermentation of sugar added to a finished still wine. These are available as red, pink, and white wines. They may be fermented in the bottle or by

a bulk process, and the gas pressure should be only that which is derivable from the actual fermentation.

Red or pink. The red sparkling wines, known as California sparkling burgundy or *champagne rouge,* are usually blended wines, and their quality is dependent upon a clean, vinous character rather than upon the varietal characteristics of any given variety. They should be a medium ruby-red; amber or tawniness is objectionable. The residual sugar content may vary between 1.5 and 5.0 per cent.

The pink sparkling wines, or pink champagne, are definitely pink and free from amber. They are fruity, fresh, and clean and usually possess no particular varietal character. The sugar ranges from 1.5 to 5.0 per cent.

White. In the class of white sparkling wines, two general types are available. One is the California sparkling muscat, which possesses a definite muscat aroma. It is of medium-to-full body and contains at least 4 per cent residual sugar. Wines of both Muscat Canelli and Muscat of Alexandria are used.

The other white sparkling wine is California champagne. Its only qualifications are that it is white and sparkling. It may or may not possess any particular varietal character; some California champagnes are made predominantly of French Colombard or Chenin blanc (often called White Pinot), a few contain Folle Blanche, and an occasional sample has some Chardonnay or Pinot noir. Champagnes are classified according to three levels of sweetness: *brut,* less than 1.5 per cent residual sugar; *demi-sec,* 1.5 to 3.5 per cent sugar; and *sweet,* more than 3.5 per cent sugar. These limits should be rigidly adhered to.

The *carbonated wines* are those to which carbon dioxide has been added under pressure. This group includes both red and white wines of no particular varietal character. In other respects they resemble the sparkling wines.

The present wine class of low carbon dioxide content finds only one representative in California, the *moscato amabile,* which is a muscat wine of low alcohol and low carbon dioxide pressure, containing some residual sugar. At present this wine pays the regular champagne tax.

The low-carbonation class (less than 7 lbs. per square inch of pressure) may include white, pink, and red wines, with or without residual sugar. The amount of carbonation will vary from a slight gassiness or prickle (*pétillance*) on the tongue to a definite evidence of bubbles in the wine.

Vermouth

The vermouths are a class of wines to which an infusion of herbs has been added. These wines may be either dry or sweet. They possess a characteristic balanced herb flavor; no single herb characteristic should predominate. The alcohol content varies from 15–19 per cent.

Dry vermouth. This is also called French-type vermouth; it is yellow and has a medium herb character. The wine should have a pronounced bitterness, and the sugar content should be less than 4 per cent, preferably 1–2 per cent.

Light dry vermouth is straw-color to light yellow (due to removal of color with charcoal), but its other characteristics are similar to those for dry vermouth.

Sweet vermouth. This type is dark amber, slightly bitter and quite sweet, the sugar content ranging from 10 to 19 per cent. The herb character is pronounced but well blended. It is best with a slight muscat aroma. Vanilla should not be noted.

Miscellaneous Types

Recently, flavored wines have appeared on the American market. These carry proprietary names. Generally they have a citrus or herb flavoring, are sweet (10 to 14 per cent sugar), and usually have the alcohol content of dessert wines.

Many fruit wines are produced in California: apple, blackberry, cherry, loganberry, raspberry, strawberry, etc. They should have a strong fruit aroma. Generally, they are very sweet and have the alcohol content of dessert wines, though for certain markets the alcohol content may be less. They should be free of added benzaldehyde or other flavoring constituent.

V. THE SENSORY EXAMINATION

Sensory testing is a costly and time-consuming undertaking. However, if worth doing at all, it should be done carefully and systematically to ensure the validity of the results. Among other considerations, proper facilities must be provided for testing, and the panel of testers must be carefully selected and trained. A statistical approach must be employed in the analysis of the results, since this is the only method by which inferences and reliable conclusions may be drawn. The design of the experiment, in particular, is of extreme importance; it should be such that all the advantages of the statistical method can be used. The best design is one that will produce the required information with the greatest degree of statistical and psychological confidence and also one that will require the fewest replications. Statistical confidence relates to a control of the errors of the first and second kind,[8] while psychological confidence deals with such factors as taste fatigue, which is due to the number of samples involved in tasting any one type of sample.

Another equally important point is the actual testing technique. Different food items require different testing procedures because of effects arising from both their chemical and physical properties. When edible oils are being tasted, the palate must be cleared between tastings to remove the cloying sensation. Pungent substances generally cause rapid sensory fatigue. Carry-over of taste sensation from sample to sample is noted with some foods. Wines, therefore, create their own sensory-testing problems. Add to this the environmental conditions and the tasters' physical condition, and it is evident that the actual testing must be conducted with extreme care.

Technique of Testing

Under no circumstances should one make a critical testing unless he is in good physical and mental condition. He should not be rushed. For reasons given previously, he should be reasonably hungry. Helm and Trolle (1946)

[8] An error of the first kind is the chance that the hypothesis of no real difference may be rejected when it is true—that is, that a difference will be found when no real difference exists. An error of the second kind is the chance that a taster may overlook a real difference.

Fig. 1. Glasses for tasting. *A*, brandy snifter, *B*, glass especially good for smelling, and *C*, general purpose tasting glass.

and Rouleau (1953) have noted the possible disadvantage of poor dental fixtures or of oral diseases on tasting ability.

Physical conditions. The analytical laboratory is *not* the proper place to conduct an organoleptic examination. The taster should be as comfortable as possible, and the room for tasting should be conducive to concentration and critical examination of the samples. It should be neither too cold nor too warm, and there should be no extraneous noise or odors. Grazzi-Soncini (1892), Mathieu (1911) and Brunet (n.d.) have especially noted that the tasting room should be quiet.

It should be painted a neutral gray or an off-white so as not to interfere with the inspection of appearance and color of the samples. (In a beer-tasting room, however, Byer and Grey [1953] used pale green.) For observing color, the illumination should be of constant and sufficient intensity. If a good northern skylight is not available, artificial light of daylight type is best. To prevent the accumulation of odors in the room adequate spittoons or sinks to carry away the expectorated wine are essential. A dentist-type spittoon is best. Unobtrusive air conditioning is very desirable. Atmospheric pressure has been suggested as another variable influencing the amount of odor and thus partially explaining the variation in results from one day to the next. It is difficult to see how this variable can be readily eliminated. Humidity may also affect odor, but it is doubtful whether the effect is important. Since humidity control is possible, however, this should perhaps be studied. Individual tasting booths are very desirable.

At the other extreme, Beattie (1949) has warned that the test conditions should not be too artificial: "Even a blindfold test creates a feeling of unreality, and may lead to queer sins of omission or commission on the part of the panel."

Tasting equipment. A proper glass for tasting is very important. Several types are shown in figure 1. A tulip-shaped glass of about 8-oz. capacity is best. A thin glass, because it can be warmed in the hands, is to be preferred. Mathieu (1911) recommends that the taster always use the same type of glass.

A silver tasting beaker (fig. 2) is valuable in observing the color of red wines but not suitable for the olfactory examination. The beaker does not permit swirling the wine and concentrating the odor in the narrowed aperture as a tulip-shaped glass does. Brunet (n.d.) reports that these are found largely in France, where they are used mainly for young wines. In the Burgundy district, however, they are used for most wines in spite of their deficiency.

Washing glasses is always a problem. Byer and Gray (1953) quote Bengtsson as recommending washing beer glasses without soap or soda, rinsing them with distilled water, and allowing them to dry without being wiped. Aside from being unsanitary, this procedure will not work with wine glasses, since red wines will gradually stain the glass. The best procedure seems to be to wash them in hot water with a detergent, rinse in warm water and in distilled water, and then allow them to dry. During drying, the glass should be hung from the stem or placed on a towel. If dried with a towel, the towel should be specially washed and thoroughly rinsed.

Temperature of the samples. Grazzi-Soncini (1892) recommends that white wines be tasted at 50° – 54° F and reds at 54° – 60° F. In our opinion these figures are too low. Brunet (n.d.) prefers 50° – 54° F for whites and 54° – 59° F for reds, but he notes that a slightly higher temperature will bring out the qualities and defects of a wine. Puls (1939) recommends the same range for whites, but 60.8° – 64.4° F for reds. Klenk (1950) says 50° – 59° F for whites and 54° – 68° F for reds is ideal. These ranges seem to be too wide. Troost and Wanner (1955) suggest 52° F for ordinary white table wines and 54° – 55° F for high-quality white table wines. For red table

Fig. 2. Silver tasting cup.

wines and for dessert wines they recommend 57° – 63° F. For sparkling wines they find 48° – 50° F best, and for spirits 55° F. We recommend a temperature of 55° – 60° F for white table and sparkling wines, 60° – 65° F for dessert wines, and 65° – 68° for red table wines.

For beer Rouleau (1953) recommended 54° – 55° F with a maximum of 59° F. However, he did not find a difference between 44° and 54° F. Byer and Gray (1953) recommended 50° F for beer but noted that Bengtsson for Swedish conditions preferred 54° – 59° F. Naturally, gustatory sensitivity decreases at low temperatures, and good and bad odors are less volatile; however, the loss of carbon dioxide is slower, and the wine (or beer) is tasted more nearly under the normal conditions of drinking. Also, what the taster is accustomed to is important. If he normally tastes wines at 65° F, it is an error to force him to taste them at 50° F and expect his critical judgment to be unimpaired. However, in panel work one should train the tasters to expect a certain temperature for each type.

Number of samples. It is obvious from what has been said about adaptation (p. 485) that tasters should not be asked to judge too many samples at a single tasting. With beer, Virden (1949) and Slater and Frenkel

(1948) considered six samples in paired tests to be the maximum for a single tasting. The tests conducted by Hinreiner *et al.* (1955*a*) have usually employed no more than this. However, in the latter case the tasters were available only once a day, and usually a single type of wine with small differences in composition or quality was being judged. If the judges were available several times during the day and if the composition or quality range was large, it would be practicable to evaluate a much larger number of samples. There is also some doubt whether a single wine can be as accurately scored as a group of similar wines of varying quality.

On the other hand, Grazzi-Soncini (1892) and Mathieu (1911) strongly recommended that wines of a similar type be judged together. The former also believed that the number of wines tasted should be regulated but gave no limit. Brunet (n.d.), however, allowed a maximum of 50 samples at a sitting, and Puls (1939), while recommending 30 to 40 per day for inexperienced tasters, said that experienced tasters may handle 80 to 100 per day. Some experimental justification for this may be found in Pfaffmann's (1954) results where fatigue was demonstrated to be less of a problem than is commonly assumed.

We feel that the number of wines to be tasted should be governed by the amount of difference between the samples, the ability of the judges, and the degree of refinement desired in the results. For best results, probably not more than five should be judged at a time. This can easily be increased to 10 if the differences are large or if some rest is given during the tasting.

Coding. Naturally, the identity of the samples must be withheld from the tasters. Some form of code marking must be used. This is best done beforehand outside the tasting room. The code should not suggest any particular order or quality to the panel. The markings should be quite random and not permit bias on the part of the taster. Usually a letter and a number are used. For example, in a paired test the pairs might be numbered C96–C53, M31–M16, F12–F23, etc. Or two-digit numbers taken from a book of random numbers may be used.

In a triangular test, because of the tendency to select the middle sample, it is probably best to arrange the glasses in a triangle on a tray before presenting the samples to the judges. The code would be similar: B12–B39–B28, K41–K73–K19, R14–R37–R17, etc. Ishler *et al.* (1954) have shown that code bias exists for codes such as A1–A2, or against certain letters.

Examination procedure. The taster should always make his observations systematically. Mathieu (1911) specifically notes that some tasters get poor results when they change the order of their observations. The best system is to observe the appearance first. This may give valuable clues as to defects. The inspection of the color may also reveal indications of defects in the wine. In some cases—for example, in buying wines or in judging them in a competition—the deficiencies in appearance or color may be so serious as to disqualify the wine and make actual tasting or even smelling unnecessary or of secondary importance. This will save the tasters unnecessary work and preserve their efficiency for other samples. Appearance and color may, in addition, alert the experienced taster to be on the lookout for certain tastes and odors. His sensitivity is thus increased (see p. 488). After visual inspection, he proceeds to olfactory examination and finally to actual

tasting. When several samples are to be compared, it is best to note the visual appearance of all the samples first and then proceed to the olfactory step.

Tasting technique. After color and appearance have been inspected, the wine is smelled. Quick, full sniffs should be taken. The wine is then swirled in the glass and sniffed again. Again the taster should proceed systematically. First he should concentrate on identifying any off-odors (p. 494), then on attempting to identify varietal aromas and, if possible, quantifying his impression. In the same manner the nature and amount of bouquet are investigated.

Finally, a small amount of wine is placed in the mouth and moved over the entire tongue so as to make contact with taste buds of all types. The taster should follow a definite sequence of investigation here: degree of acidity, dryness or sweetness, and astringency. Since the taste buds for astringency are at the back of the tongue, the perception of tannin may be delayed. Some tasters throw back the head and almost gargle the wine the better to observe its astringency. Others "chew" the wine, while still others allow it to warm up in the mouth. Experienced tasters also draw some air through the wine in the mouth to churn the wine and thus volatilize additional odors. These are then smelled as the air is exhaled.

When many samples are to be tasted, it is essential to expectorate the wine. However, if there are only two or three wines, some of each may be swallowed. The "finish" of a wine is probably largely a matter of astringency, and in swallowing the wine the taster is assured that it reaches all the taste buds sensitive to bitterness. Rinsing the mouth with water between every three or four samples is common. If the tasting is prolonged, some tasters eat a little bread or cracker to allay hunger and to remove taste effects (Brunet, 1908; Baten, 1947).

Order of tasting. The usual order of tasting when there is more than one type of wine to be judged at a sitting is from low to high alcohol, dry to sweet, and white to red. Grazzi-Soncini (1892) and Mathieu (1911) suggest starting with the weakest (lowest alcohol), thinnest, or greenest wines, reserving the maturer, more aromatic, smoother, and high-alcohol wines for the last. Brunet (n.d.) suggests starting with the poorest, driest, and lightest wines and then proceeding to the more alcoholic and richer (sweeter, more body, etc.) wines. He says that dry and sweet and red and white wines should never be mixed. With this we agree. In comparative judgings, where a definite order of tasting is used, the arrangement in serving order within the class must be randomized for each taster.

Influence of one sample on another. The effect of wines of varying quality on each other is well known. However, it is seldom considered in the design of tastings. So far as the authors know, all judgings at fairs and similar events suffer from this defect. Each judge receives the samples in the same order. It would be preferable for the order to vary from judge to judge so that a poor grade of wine could equally well precede or follow a better one. The effect is somewhat minimized at the fair judgings because there is some tendency for the judges to compare the samples back and forth without regard to order. With a panel of expert judges, tasting five wines in blocks of three and four samples, Filipello (1957) found that the order

of serving had no effect when a quality rating scale was used. In certain laboratory panel tests, however, the order must be randomized. Byer and Gray (1953) demonstrated with beer that the preferences of the tasters depended upon the order in which the samples were tasted. Mathieu (1911) stated that a strong odor of one sample would tend to reduce even further the lesser odor of a following sample.

Time of day. While Goetzl's (1949) work suggests that tasters should be hungry when they are judging, no critical experiments seem to have been made with wines. We recommend that the time of tasting be standardized as much as possible and believe the period between 10 and 12 in the morning is likely to prove best.

Number of judges. The number of tests and the number of judges needed depends upon the objective of the tests. For difference testing a smaller but more expert panel will give better results. Baker *et al.* (1954) showed that to detect organoleptic differences, if they exist, it is important to use the best tasters available. A combination of tasters of varying ability gives results that are much poorer than the best but somewhat better than the worst. While it is conceivable that a single expert (good taster with low variability) might afford economies of replication, time, and cost, a panel of more tasters will usually produce the most reliable results. Mathieu (1911) strongly recommended experienced panels rather than single experts.

Obviously we do not always know the relative ability of the tasters. Furthermore, the number of samples that a single judge can taste efficiently is limited. Finally, there are limitations on the number of judges that can be obtained and on the time they can give. For routine winery work we suggest expert panels of five to ten tasters with duplicated tastings. Krum (1955) suggests ten to thirty judges for routine difference tests.

Wholly unrelated to this problem is that of consumer-preference tests. In this case, much larger groups are needed, randomly chosen, in order to obtain an average group response. In mass-panel wine-preference tests at California fairs, a reproducible mean value of the degree of liking could be obtained for any wine from approximately 100 judgments for any group investigated. Since there were three age groups and three frequency-of-wine-use categories, there were nine subcategories, and it was necessary for 900 people to taste any one wine in order to obtain full information concerning it. See Anon. (1958) for further data on consumer panel size.

Training the panel. For best results it is absolutely essential that the panel be trained for the type of difference it is likely to encounter. Even experienced judges need to have their memory refreshed by practicing on samples with known differences in the range likely to be found. Of course, in some cases we do not know whether or not a difference exists.

When a known chemical difference is being studied, as with sugar, acid, and sulfur dioxide, the panel should practice for several days with wines made up to a range of differences. If the odor is not known—that is, is not chemically identifiable, for example, as filter pad—a dilution test in which the strongest odor is diluted with a neutral wine of identical composition can be used to train the panel.

It may be found that judges capable of detecting small differences in table wines will be less sensitive with dessert wines or brandies. Some system

of checking performance is therefore needed. The best procedure is to prepare samples with known small differences of the same type to be expected in the test and to give each member of the prospective panel a paired (p. 523), triangular (p. 526), or sequential (p. 530) type of test. Those who do not find the difference can then be eliminated from the final panel. Obviously, only those whose employment will permit them to participate regularly in the tasting should be considered for a panel.

Points to Remember

1. All the samples of a particular lot must be identical. This means that if the wine is to be served to a number of tasters, all the wine should come from the same bottle. If more than one bottle is required, the necessary number of bottles should be decanted together before the samples are taken. When wines are to be served at different times, this procedure is impractical, since the wine will gradually change in quality after decanting. In this case, one should select bottles that are adjacent in the bin or bottles that are known to have had identical treatment. If a single sample is taken, it must be representative of the whole lot. This is by no means as simple as it seems, particularly with older wines. Even with a newly bottled lot, care must be taken not to sample the very first wine bottled (which is usually more aerated and higher in metals than later, more typical samples). Even sampling from a tank is subject to error if the sample is taken with an unclean thief, if it is taken only from the top or the bottom, or if the contents of the tank have been disturbed for one reason or another. Clean glass thiefs are preferred. If samples are taken through a cock on the side of a tank, enough liquid should run through to clean out any high-metal wine in contact with the metal.

2. While the matter of fatigue is still controversial, it is better to limit the number of samples at one sitting. With products of pronounced flavor, such as vermouth, Beattie's (1949) results with ginger beer suggest a lengthy rest after even a pair of samples.

3. The procedure should be standardized to the most minute detail. This means that temperature, method of pouring, size and shape of glass, number of samples, etc., should be the same for all judges. For further details concerning tasting see Bengtsson and Helm (1946), Boggs and Hanson (1949), Helm and Trolle (1946), Mathieu (1911), Peryam (1950), Peryam and Swartz (1950), and Wood (1949).

VI. SYSTEMS OF TESTING

Difference Tests Involving Simple Difference, Difference in a Given Constituent, or Difference in Preference (i.e., Quality Difference)

"Is there a difference between the wines?" is the only question asked in simple difference tests. In such tests it is assumed that the two wines being tasted are different. Such tests are used to detect differences between samples of different blends, samples treated in different ways, wines of the same variety from different regions, wines made from grapes harvested at different times, or between wines fermented by different procedures, bottled

Approximate colors of California white wines. Upper left, chablis; upper right, muscatel; lower left, dry sherry; lower right, sweet sherry.

Approximate colors of California red table wines. Upper left, *rosé;* upper right, claret; lower left, burgundy; lower right, Cabernet or young red wine.

at different ages, or stored under different conditions. Difference tests are also employed to determine the minimum detectable difference in some constituent, such as sugar or alcohol. In all such tests all other constituents should remain the same. *It is obvious that if no difference can be established, the subsequent question of preference between the samples is not necessary.*

In analyzing statistically the results of taste tests, simple and constituent testing must be distinguished from preference testing. In difference testing the taster may be asked to determine if a particular wine has more (or less) of some constituent, a condition that can usually be determined by a specific chemical test, than another wine used as a standard. If the first wine has more of the constituent than the standard and if the taster so indicates more often than might be expected by chance, the experimenter can infer that the taster does possess some ability to detect the particular constituent. This is known as a one-sided, or "one-tailed," test, since the taster shows ability only if he can make the proper choice an unusually large number of times.

In the case of preference or quality testing, if in a series of tests a taster is asked to state a preference between two wines, a large number of selections of one wine over the other may indicate, on the basis of the taster's judgment, a significant quality or preference difference between the two wines. Since either wine may be the preferred one and since the selection of the one wine an unusually small number of times is just as interesting as its selection an unusually large number of times, the appropriate test is two-sided, or "two-tailed." In the examples to follow the appropriateness of one- and two-tailed tests will be indicated.

The results of an experiment in taste testing, however, have little meaning unless the judges demonstrate ability to detect differences in taste when such differences do exist, and do so with good consistency. In many cases these differences are small and are not easily detected. It is therefore apparent that in such tests the panels should consist of judges of the greatest degree of sensitivity. Baker *et al.* (1954) particularly emphasized the advantage of using the most skilled judges for sensory tests.

The samples to be tested should not differ in color unless color is the difference factor, since color differences may make it possible for the panel to identify the samples without tasting or smelling them. Even if the judging system is designed to prevent identification of a color difference (by using interchangeably two or more series of wines of the same general type and similar color differences), color difference should be avoided. Prejudices for or against a certain color may impair the reliability of the judges. Color or appearance differentiations can be avoided, however, by serving the wines in glasses, the exteriors of which are painted black. Reducing the light in the tasting room will also decrease the possibility of observing slight color differences. The judges should not be blindfolded, however, as this device reduces their olfactory acuity (and possibly even their taste acuity).

Paired-sample test. In this test the taster is given two samples and is asked to state which he prefers, or to pick out the sample higher in some constituent (see Cover, 1936). This procedure may be carried out by one taster several times or by a panel of tasters one or more times. In the several trials the wines are presented in identical glasses but in different orders.

On the basis of the null hypothesis—that is, that there is no difference between the wines—the probability p of a taster identifying by chance a particular wine in each of the several trials is ½.

In paired-sample tests the corrected χ^2 distribution is simple to apply to the results in determining differences between samples or in selecting the best judges for a panel. To analyze the results of such a test, let N equal the total number of tests, X_1 the number of tests favorable to wine No. 1, and X_2 the number of tests favorable to wine No. 2. Then

$$\chi^2 = \frac{(|X_1 - X_2| - 1)^2}{N}$$

where $| X_1 - X_2 |$ indicates the absolute value or the positive value of the difference.

When only difference (one-tailed) identification is requested, the calculated value of χ^2 must equal or exceed 2.71 for significance at the 5 per cent level ($p = 0.05$); 5.41 for significance at the 1 per cent level ($p = 0.01$); and 9.55 for significance at the 0.1 per cent level ($p = 0.001$). When N is less than 5, the exact probabilities should be calculated by the binominal theorem.

For example, a taster is given in one glass a dry white table wine of low volatile acidity and in a second glass the same wine to which a detectable and undesirable amount of ethyl acetate has been added. Eleven times in 16 trials he identifies correctly the sample to which the ethyl acetate has been added. In this case $\chi^2 = (| 11 - 5 | - 1)^2 / 16 = 1.6$. Since this value is less than 2.71, this taster, at the 5 per cent level of significance, would not be an acceptable judge of this constituent in this wine.

When comparisons are made on the basis of quality preférence (two-tailed test), the calculated value of χ^2 necessary for significance at the three levels must equal or exceed 3.84, 6.64, and 10.83 respectively. For example, if another taster made the comparison 20 times and preferred wine No. 1 in 15 of the trials, then $\chi^2 = (| 15 - 5 | - 1)^2 / 20 = 4.0$. This indicates that on the basis of this taster's performance wine No. 1 is significantly better than wine No. 2, and since 4.0 exceeds 3.84 only slightly, the odds that this result would occur by chance are slightly less than 1 in 20 (5 per cent). For another taster who preferred wine No. 1 12 times out of 20 trials, $\chi^2 = (| 12 - 8 | - 1)^2 / 20 = 0.45$. On the basis of this small value of χ^2, one can say that this judge has no significant difference in preference between the wines. If, however, a definite and significant difference in quality is known to exist, then this taster is shown not to be an acceptable judge of the difference in quality between the two wines.

It is interesting to note that if the results of the two tasters in the above example are combined, $\chi^2 = (| 27 - 13 | - 1)^2 / 40 = 4.2$, which is also significant at the 5 per cent level and indicates a real quality difference between the two samples.

The number of correct judgments (one-tailed test) or agreeing judgments (two-tailed test) necessary for significance in each case can be determined from table 2. From the table the number of agreeing quality

TABLE 2

SIGNIFICANCE IN PAIRED TASTE TESTS ($p = \frac{1}{2}$)

Number of tasters or tastings	Minimum correct judgments to establish significant differentiation (one-tailed test) Probability level*			Minimum agreeing judgments necessary to establish significant preference (two-tailed test) Probability level*		
	.05	.01	.001	.05	.01	.001
7	7	7	..	7	..	..
8	7	8	..	8	8	..
9	8	9	..	8	9	..
10	9	10	10	9	10	..
11	9	10	11	10	11	11
12	10	11	12	10	11	12
13	10	12	13	11	12	13
14	11	12	13	12	13	14
15	12	13	14	12	13	14
16	12	14	15	13	14	15
17	13	14	16	13	15	16
18	13	15	16	14	15	17
19	14	15	17	15	16	17
20	15	16	18	15	17	18
21	15	17	18	16	17	19
22	16	17	19	17	18	19
23	16	18	20	17	19	20
24	17	19	20	18	19	21
25	18	19	21	18	20	21
30	20	22	24	21	23	25
35	23	25	27	24	26	28
40	26	28	31	27	29	31
45	29	31	34	30	32	34
50	32	34	37	33	35	37
60	37	40	43	39	41	44
70	43	46	49	44	47	50
80	48	51	55	50	52	56
90	54	57	61	55	58	61
100	59	63	66	61	64	67

* $P = .05$ indicates that the odds are only 1 in 20 that this result is due to chance; $p = .01$ indicates a chance of only 1 in 100; and $p = .001$, 1 in 1,000.

judgments (two-tailed test) in 40 trials for significance at the 5 per cent level is seen to be 27. This means that if, in 40 trials, a particular wine is preferred 27 times, one may state that the preferred wine is significantly better, and the chances of being wrong in this conclusion are only 1 in 20 ($p = 0.05$).

In the acuity taste test involving ethyl acetate (one-tailed test) 12 correct identifications in 16 trials are necessary for significance at the 5 per cent level.

Duo-trio test. In a duo-trio test a taster is presented with three samples, one of which is labeled as to identity and is often designated as the *reference* sample. Of the other two, in practice generally marked No. 1 and No. 2, one is identical with the labeled sample and the other is different. The taster is asked to choose the one he believes to be different from the labeled sample, so he simply selects either No. 1 or No. 2. As in the case of the paired-sample

test, the probability that in a single trial the odd sample will be chosen by chance is one-half ($p = 1/2$). The χ^2 test described in the paired-sample test is applicable.

For example, each of 15 testers performs a test two times with the following results:

Taster	Number of correct identifications	Taster	Number of correct identifications
1	2	9	2
2	1	10	2
3	2	11	2
4	2	12	1
5	2	13	0
6	0	14	2
7	1	15	2
8	2		23

For these data $\chi^2 = (|\ 23 - 7\ | - 1)^2\ /\ 30 = 7.5$. Since this is a difference test (one-tailed) and the calculated value of χ^2 exceeds 2.71, the tasters have been able to establish significant difference between the two samples.

If, in addition to identifying the samples correctly, the taster is asked to state which of the two wines has more of some constituent (one-tailed test) or which he prefers (two-tailed test), on the basis of the null hypothesis, the probability in any single trial of a correct judgment or of his agreeing with a previous quality judgment is one-fourth ($p = 1/2 \times 1/2 = 1/4$), and the appropriate expression for χ^2 is

$$\chi^2 = \frac{(|4X_1 - N| - 2)^2}{3N}$$

where, as before, X_1 is the number of tests favorable to wine No. 1 following correct identification. The values of χ^2 at the three significance levels are the same as previously indicated in the paired-sample test. The number of correct judgments (one-tailed) or agreeing judgments (two-tailed) necessary for significance in each case can be determined from table 3.

There is one limitation in the use of the table. In most cases the number of agreeing judgments necessary to establish significant preference is less than half the number of tasters or tastings. It is therefore possible, but not probable, that each of two wines might be selected a sufficiently large number of times to indicate for each significant superiority. Obviously, on the basis of such results significant preference is not established. In such a situation the selections of all judges having first established their ability to identify the wines properly should be analyzed as a paired test, and significance would, therefore, follow the pattern indicated in table 2, the number of tasters or tastings being those whose ability to identify wines properly has been indicated.

Triangular test. In the triangular system (Amerine, 1948*a*) three glasses, A, B, and C, are presented to the taster at the same time. Two of the glasses contain the same wine and one a different wine, and the taster is so informed. The taster is required, after tasting, to indicate whether A, B, or C is the

TABLE 3

SIGNIFICANCE IN DUO-TRIO TASTE TESTS ($p = ½$)

Number of tasters or tastings	Minimum correct judgments to establish significant identification (one-tailed test)			Minimum agreeing judgments necessary to establish significant preference (two-tailed test)		
	Probability level			Probability level		
	.05	.01	.001	.05	.01	.001
5	4	5	5	4	5	..
6	4	5	6	5	5	6
7	5	6	7	5	6	7
8	5	6	7	6	6	7
9	5	6	8	6	7	8
10	6	7	8	6	7	8
11	6	7	9	7	8	9
12	7	8	9	7	8	9
13	7	8	9	7	9	10
14	7	9	10	8	9	10
15	8	9	10	8	9	11
16	8	9	11	9	10	11
17	8	10	11	9	10	12
18	9	10	12	9	11	12
19	9	10	12	10	11	12
20	9	11	12	10	11	13
21	10	11	13	10	12	13
22	10	11	13	11	12	14
23	10	12	14	11	12	14
24	11	12	14	11	13	15
25	11	13	14	12	13	15
30	13	14	16	13	15	17
35	14	16	18	15	17	19
40	16	18	20	17	18	21
45	17	19	22	18	20	22
50	18	20	23	19	21	24
60	20	23	26	21	24	27
70	22	26	30	24	27	30
80	25	29	33	27	30	34
90	28	32	36	30	33	37
100	30	35	39	32	36	40

odd sample. He may *not* say that all are the same or all different. If the two wines are designated No. 1 and No. 2, it is obvious that they can be poured into glasses A, B, and C in the following orders: 1–1–2, 1–2–1, 2–1–1, 2–2–1, 2–1–2, or 1–2–2. By chance alone the taster would be expected to identify the odd sample correctly in one third of the trials ($p = 1/3$), and the value of χ^2 (see Roessler *et al.*, 1948) may be calculated from

$$\chi^2 = \frac{(|4X_1 - 2X_2| - 3)^2}{8N}.$$

If, in addition to identifying the odd sample correctly, the taster is asked to state which of the two wines has more of some constituent (one-tailed) or which he prefers (two-tailed), on the basis of the null hypothesis the

TABLE 4

SIGNIFICANCE IN TRIANGULAR TASTE TESTS

Number of tasters or tastings	Minimum correct judgments to establish significant differentiation ($p = 1/3$) (one-tailed test)			Minimum correct judgments to establish significant identification ($p = 1/6$) (one-tailed test)			Minimum agreeing judgments to establish significant preference ($p = 1/6$) (two-tailed test)		
	Probability level			Probability level			Probability level		
	.05	.01	.001	.05	.01	.001	.05	.01	.001
5	4	5	5	3	4	5	4	4	5
6	5	6	6	4	4	5	4	5	6
7	5	6	7	4	5	6	4	5	6
8	6	7	8	4	5	6	5	5	6
9	6	7	8	4	5	7	5	6	7
10	7	8	9	5	6	7	5	6	7
11	7	8	9	5	6	7	5	6	8
12	8	9	10	5	6	8	6	7	8
13	8	9	10	5	7	8	6	7	8
14	9	10	11	6	7	8	6	7	9
15	9	10	12	6	7	9	7	8	9
16	10	11	12	6	7	9	7	8	9
17	10	11	13	7	8	9	7	8	10
18	10	12	13	7	8	10	7	9	10
19	11	12	14	7	8	10	8	9	10
20	11	13	14	7	9	10	8	9	11
21	12	13	15	7	9	10	8	9	11
22	12	14	15	8	9	11	8	10	11
23	13	14	16	8	9	11	9	10	12
24	13	14	16	8	10	11	9	10	12
25	13	15	17	8	10	12	9	10	12
30	16	17	19	10	11	13	10	12	14
35	18	19	21	11	12	14	11	13	15
40	20	22	24	12	14	16	13	14	16
45	22	24	26	13	15	17	14	15	18
50	24	26	28	14	16	18	15	16	19
60	28	30	33	16	18	21	17	19	21
70	32	34	37	18	21	23	19	21	24
80	35	38	41	20	23	26	21	24	26
90	39	42	45	22	25	28	23	26	29
100	43	46	49	24	27	30	26	28	31

probability in any single trial of a correct judgment or of his agreeing with a previous preference (quality judgment) is one-sixth ($p = \frac{1}{3} \times \frac{1}{2} = \frac{1}{6}$), and the appropriate χ^2 is

$$\chi^2 = \frac{(|6X_1 - N| - 3)^2}{5N}.$$

Table 4 indicates how many times a taster must correctly find a difference (one-tailed test), identify (one-tailed test), or agree in a preference (two-tailed test) for or against one of the two wines to be significant. The same limitation applies in the use of this table as in table 3 where more than one wine exceeds the number of agreeing judgments to establish significance.

Triangular tests are widely used to select taste panels. The procedure has the advantage of simplicity. Obviously a taster obtaining 10 out of 10

correct scores will make a better judge than one obtaining 5 out of 10 correct scores. It is now generally recognized that people vary in their ability to detect certain taste differences. One taster may be a good judge of Cabernet odor but fail to detect slight differences in fusel oil. Hence panels should be selected for particular purposes. However, many tasters can be used for a wide variety of panels. While the triangular taste test may establish a significant difference between two wines, it may or may not prove that one is better. Tasters are usually less successful in specifying consistently which wine they prefer than in detecting differences by a proper selection of the odd sample.

Disadvantages of the triangular test are the time required to set up the test, the possibility of mixing the samples during coding, the number of tests required for significance, and the necessity of preventing the detection of differences solely on the basis of color. This is best done by serving the wines in beakers, the entire surface of which has been painted black. It has also been pointed out by Harrison and Elder (1950) and by Harries (1956) that there is an unexplained psychological tendency to select the middle sample as the "odd" one. To prevent this it has been recommended that the identification of the glasses be concealed from the taster until he has made his decision. It has also been demonstrated that the triangular test is slightly less sensitive than the paired test in revealing small differences (Byer and Abrams, 1953; Gridgeman 1955; and Filipello, 1956*a*).

More complicated systems may be used. If four glasses, two of wine No. 1 and two of wine No. 2, are presented, they can be arranged in six different ways: 1–1–2–2, 1–2–2–1, 2–1–1–2, 1–2–1–2, 2–1–2–1, and 2–2–1–1. By chance alone the taster can be expected to pair the samples correctly in one third of the cases, and the analysis of the results is identical with that of the triangular test. The probabilities of correct arrangement by chance alone for combinations of one to six samples of each wine are as follows (Lockhart, 1951):

Number of samples of wine No. 1	Number of samples of wine No. 2					
	1	2	3	4	5	6
1	1	1/3	1/4	1/5	1/6	1/7
2		1/3	1/10	1/15	1/21	1/28
3			1/10	1/35	1/56	1/84
4				1/35	1/126	1/210
5					1/126	1/462
6						1/462

The probabilities shown in the table are based on completely correct separation of the wines into two groups each containing similar samples. A taster may not make a completely correct separation but may still show evidence of ability to differentiate the samples. If 1–1–1 is considered as two correct pairs and 1–1–1–1 as three correct pairs, then in the case of four wines with three samples each, perfect differentiation would be designated as 1–1–1, 2–2–2, 3–3–3, 4–4–4 and would represent 8 correct pair-

ings; four correct pairings occur in the cases: 1–1–2, 2–2–3, 3–3–4, 4–4–1 and 1–1–1, 2–2–4, 3–3–4, 2–3–4. In the following table are shown the minimum number of correct pairings for significance at the 5 per cent and 1 per cent levels in the case of three, four, five, and six wines with various numbers of samples of each:

Number of wines	Number of samples of each	Minimum number of correct pairs for significance	
		5%	1%
3	2	..	..
4	2	4	4
5	2	3	5
6	2	3	4
3	3	6	6
4	3	5	6
5	3	5	6
6	3	4	5
3	4	7	9

For example, three different wines, X, Y, and Z, are each poured into three glasses, and the nine glasses are presented to the taster in the random order X-Y-X-Z-Z-Y-Y-X-Z. The taster is asked to identify the three wines. If his final separation is XXY, YYX, and ZZZ, he has four correct pairs. From the above table for three wines with three samples of each, six correct pairs are required for significance at both the 5 and 1 per cent levels. The taster, therefore, does not indicate ability to differentiate between the wines at these levels of significance.

Using five samples, two of one kind and three of another, arranged in a row, Harries (1956) has shown that there is a tendency to pick samples No. 1 and No. 5 as pairs. He found that if the samples were arranged in a circle this error was avoided.

Sequential analysis for selecting judges. The duo-trio and triangular tests are frequently used to determine an individual's ability as a judge. In order to show with reasonable confidence that he is doing more than guessing, the prospective judge is required to repeat the experiment successfully a considerable number of times. There is some question regarding the quality of the judges so obtained and the number of tests that should be given. In general, too little testing is done because of the time involved and limitations of suitable experimental material.

Sequential methods afford a saving of time and material over most other selective procedures; they are equally reliable and are well suited for the selection of judges. In a sequential testing plan the number of observations is not predetermined, and the decision to terminate the experiment depends, at any stage, on the results of previous observations. In taste testing a small number of trials is desirable, since taste and olfactory fatigue accompanies repetition and only a few trials can be performed at one sitting. Methods that minimize the time interval for an individual to demonstrate

discriminating ability and that reduce the cost of preparation and performance of the experiment are to be desired.

Systems of sequential analysis have been developed by Wald (1947) and Rao (1950). Such systems involve the formulation of a rule by which one of the following decisions can be made at any stage of the experiment: (1) accept the taster as a judge, (2) reject the taster, (3) continue the experiment by taking an additional observation.

The procedure developed by Wald is as follows. Let p be the true proportion of incorrect decisions in paired, duo-trio, or triangular tests if the judge could continue testing indefinitely. Then $q = 1 - p$, the proportion of correct decisions, may be defined as the judge's inherent ability under the test administered. Values $q_0 = 1 - p_0$ and $q_1 = 1 - p_1$ can be so specified that individuals having abilities equal to or greater than q_0 will be selected for the taste-testing panel and those with abilities equal to or less than $q_1 = 1 - p_1$ will be ruled unacceptable.

The criterion for accepting or rejecting potential judges is given by two parallel straight lines, lines L_0 and L_1, which are uniquely determined by assigned values of p_0, p_1, α, and β, where α is defined as the probability of rejecting an acceptable judge and β is the probability of selecting an unacceptable judge. These are the well-known errors of the first and second kind (see p. 515). If potential judges are in good supply, α may be selected large and β small. The lines L_0 and L_1 divide the plane into three regions, one of acceptance, one of rejection, and one of indecision. The equations of these lines may be written

$$d_0 = a_0 + bn \text{ (lower line)}$$

$$d_1 = a_1 + bn \text{ (upper line)}$$

where n represents the total number of trials, d the number of incorrect decisions, b the slope, and a_0 and a_1 the intercepts on the vertical axis.

The slope and intercepts are given by the following formulas:

$$b = \frac{\log \dfrac{1 - p_o}{1 - p_1}}{\log \dfrac{p_1}{p_o} - \log \dfrac{1 - p_1}{1 - p_o}}$$

$$a_o = \frac{\log \dfrac{\beta}{1 - \alpha}}{\log \dfrac{p_1}{p_o} - \log \dfrac{1 - p_1}{1 - p_o}}$$

and

$$a_1 = \frac{\log \dfrac{1 - \beta}{\alpha}}{\log \dfrac{p_1}{p_o} - \log \dfrac{1 - p_1}{1 - p_o}}$$

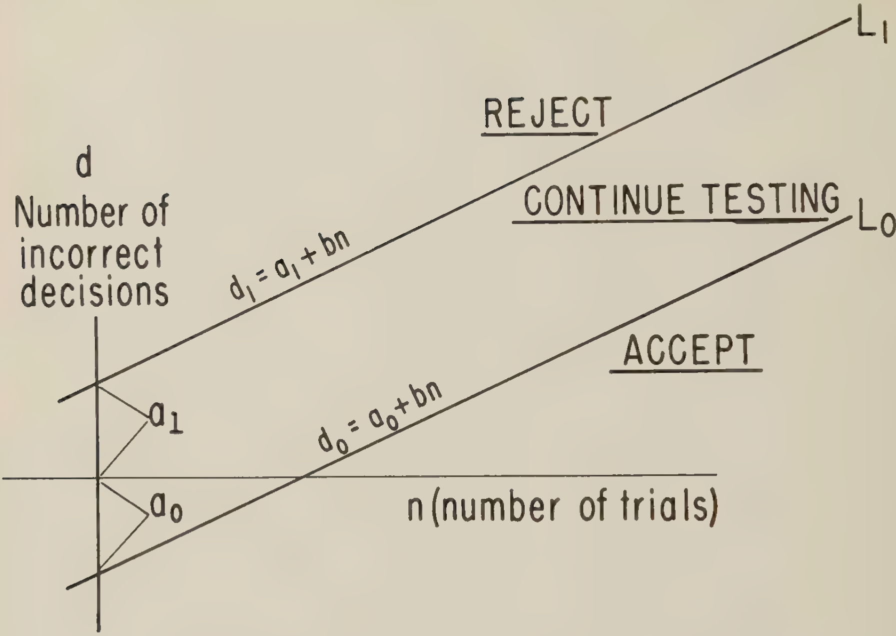

Fig. 3. Graphical representation of sequential analyses.

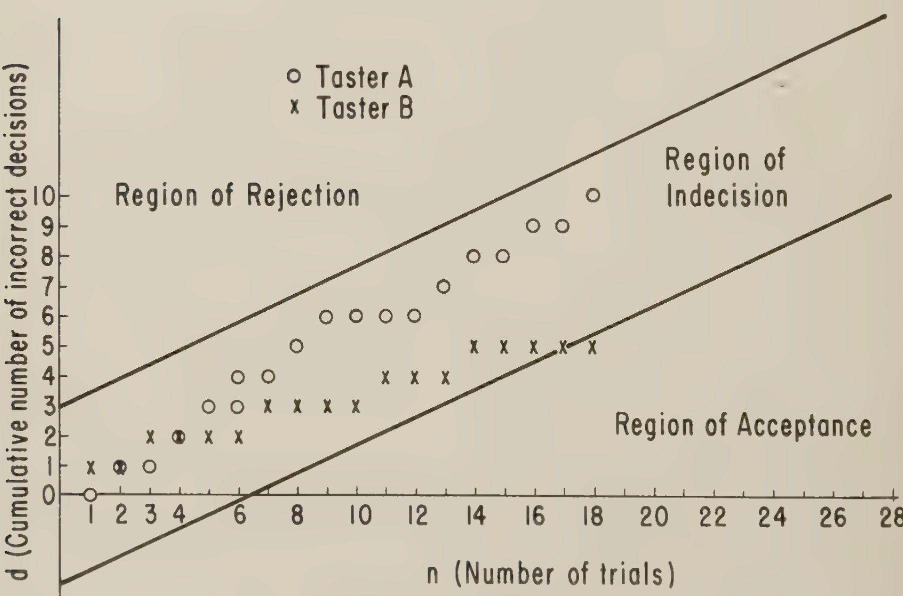

Fig. 4. Cumulative results of two tasters.

The way in which the plane is divided by the lines L_0 and L_1 is shown in figure 3.

After each taste trial the experimenter plots on the graph the point representing the total number of incorrect decisions against the total number of trials. Each plotted point is therefore one space to the right of the preceding point or one space higher, depending on whether the judge made a correct or an incorrect decision in the last trial. Testing continues until a plotted point falls on or above the upper line, or on or below the lower line. In the former case the judge is rejected, and in the latter he is accepted as a judge.

TABLE 5

RESULTS OF TASTE TESTS WITH TWO TASTERS

Number of trials (n)	A		B	
	Wrong decisions	Cumulative wrong decisions	Wrong decisions	Cumulative wrong decisions
1	0	0	1	1
2	1	1	0	1
3	0	1	1	2
4	1	2	0	2
5	1	3	0	2
6	1	4	0	2
7	0	4	1	3
8	1	5	0	3
9	1	6	0	3
10	0	6	0	3
11	0	6	1	4
12	0	6	0	4
13	1	7	0	4
14	1	8	1	5
15	0	8	0	5
16	1	9	0	5
17	0	9	0	5
18	1	10	0	5

For example, suppose that a triangle test is being used as a basis for selecting judges in a sequential procedure. Reasonable ability limits are $q_0 = 1 - p_0 = 0.65$ and $q_1 = 1 - p_1 = 0.40$; that is, individuals with abilities 0.65 or greater will be accepted and those with abilities 0.40 or less will be excluded as judges. Select $\alpha = \beta = 0.05$; that is, the probabilities of rejecting a satisfactory judge and accepting an unsatisfactory one are equal.

Substituting these values of p_0, p_1, α, and β in the equations for b, a_0, and a_1, one finds that

$$b = 0.474, \quad a_0 = -2.87, \quad \text{and} \quad a_1 = +2.87$$

The equation of L_0 is $d = -2.87 + 0.474n$; and of L_1, $d = 2.87 + 0.474n$.

The results of two tasters are given in table 5, where 1 indicates a wrong judgment and 0 denotes a correct one.

In figure 4 the number of trials and the cumulative number of incorrect decisions are plotted as points. After 18 trials taster B may be accepted as

a judge, but the points representing taster A after this number of trials are still in the region of indecision. He will have to continue tasting to qualify as a judge.

Various values of p_0, p_1, α, and β may be used in selecting judges for an experimental panel. The experimenter may believe that more stringent limits than those used in the illustrative example are appropriate. As p_1 approches p_0 the number of trials required is increased. One means of reducing the number of trials required is to increase α, accepting a greater risk of rejecting a capable judge. This may be done, for example, if large numbers of individuals are available as prospective judges so that the rejection of an able one is not regarded as a serious error.

If many wines are to be compared by several judges, the experiment may be designed as a randomized complete block or a randomized incomplete block and the results analyzed by the analysis of variance procedure. These techniques are explained in numerous publications and an account of them will not be given here.

Ranking by Scores

Ranking by scores means giving the wine a numerical rating proportionate to its quality. Most systems based on ranking by scores have no rational basis because the scoring systems proposed are based on some arbitrary scale of values, 0–100, 0–50, etc. Whether or not the judges have a clear concept of the quality associated with each of the steps in the scale, or even if all the judges have the same concept, is not stated. The advantage of score cards is that a number of samples can be compared. Score cards systematize the tasters' evaluation of the wine and insure that no single characteristic is omitted. Moreover, score cards can be very simple (as hedonic score cards are) or as complete as may be desired. A complete score card provides a permanent record of the taster's impression of the wine on a certain date. This is possible by using a quantitative notation on the score card or by assigning different numerical values for different degrees of odor or taste. Furthermore, if score cards are used tests made at different times can, to a certain extent, be compared with each other.

There are, however, several distinct disadvantages to score cards in addition to the theoretical ones mentioned above. These have to do with the possible injustice of partial additive scores for different characteristics adding up to a figure which is not a true measure of quality. Consider a score card for five characteristics. Wine A scores 5–5–5–5–0 for a total of 20, whereas wine B scores 4–4–4–4–4 for a total of 20. In spite of this, wine B is probably superior to A. It has been suggested that if a wine scores 0 or below some given value for any single characteristic, it be arbitrarily assigned a certain maximum score. The use of weighted geometric or harmonic means has also been suggested.

The usual wine score card quite rationally divides the score into clarity, color, smell, taste, and general quality. But who is to say that, of a total score, color should have a certain percentage value for all types of wine? Or even for wines of the same type? It is not usually known whether the difference between a total score of 10 and 12 is the same to the taster as the difference between 24 and 26 or 84 and 86.

TABLE 6

SUGGESTED WINE SCORE CARDS

	Blaha (1948)	Gallay and Benvegnin (1950)	Prostoserdov (1948)	Paleiri and Garoglio (1947)	Cruess (1935)			Hallgarten (1951)	Amerine and Joslyn (1951)	Amerine (1958)	TAC*
					Dry red	Dry white	Dessert				
Appearance	12	5	20	20	15	20	10	3	10	2	2
Color	12		20	20	10	10	10	3	10	2	
Bouquet	12	5	20	20	15	10	15	9	15	4	4
Volatile acid	..	..	..	..	15	15	10	..	10	2	..
Flavor or taste	20	5	20	40	15	20	15	..	15	2	4
Acidity	..	..	..	..	10	10	10	..	8	2	..
Sugar	..	..	..	..	10	10	20	3	8	1	..
Body	2	5	..	..	..	..	..	9	6	1	..
Alcohol	.	5	..	..	..	..	..	3	..	..	..
Astringency	..	..	..	..	10	5	..	..	10	2	..
Finesse	.	5		.	.	..	.	3	.	.	..
Freshness of taste		5			..	..	.	.			..
Age	.	..						3	.		..
Harmony	..	5	..	..	..	..	..	..	..	..	..
Varietal character	20	..	20	..	..	..	..	..	..	..	..
General quality	12	..	..	..	..	..	10†	..	8	2	..
Maximum	90	40	100	100	100	100	100	36	100	20	10

* Unpublished score card of the Technical Advisory Committee of the Wine Institute, San Francisco.
† Rancio or "sweet wine" taste.

Marescalchi (1949) questioned 20 Italian wine experts as to what percentage of a 50-point score card should be allocated to alcohol, acidity, extract, color, appearance, etc. There was little agreement between the experts. However, with consultation many of the disagreements in interpretation might have been adjudicated. Kielhöfer (1949) has also noted the disagreement of authorities as to the relative number of points that should be assigned to each of the sensory factors. Mensio (1957) uses a 30-point system.

A number of score cards are shown in table 6. As far as we know, none of them have been based on sound statistical analyses of the tasting data.

Hinreiner (1956) has suggested that better results might be obtained by having each judge classify the samples for only the characteristic he is most qualified to judge. Thus, one judge might rank or grade the sample for appearance, another for color, another for off-odors, etc. By this system a limited number of judges can handle a larger number of samples than if each had to score each sample for all characteristics. This system really amounts to the single-expert system. If all the judges are competent for the characteristic they are judging, the results should be very good. If, however, the judge for desirable odor has an off-day or is less competent, his results will have a disproportionate effect on the final scores and they will be less reliable. This raises the problem of measuring a judge's performance. Boggs and Hanson (1949) preferred using the correlation coefficient for the taster's first score and his duplicate score for a series of samples of *varying* quality because it relates reproducibility to the magnitude of differences shown for unlike samples. When they found unsatisfactory performance, all the judge's results were discarded. It seems to us that a better procedure would be to do more thorough panel selection according to the objectives of the test. Elimination of one judge's results does not seem a wholly consistent objective practice.

It is certain that score cards with more than 20 steps are somewhat artificial, since even experienced tasters cannot differentiate a greater range of quality than 20 steps. A 20-step score card was proposed by Brunet (n.d.). Probably for inexperienced judges a scale of only 10 steps is sufficient. We do not recommend that any scale should have more than 20 steps.

The score card used in this laboratory was originally based on a 100-point scale, but it was soon recognized that a range of only 20 points—65 to 84—was utilized. The full description of the 20-point score sheet is given below. In this form it has some of the aspects of the flavor profile used by Cairncross and Sjostrom (1950). It has been observed that experienced tasters tend to give an over-all score directly without scoring on each characteristic. However, until complete familiarity with the score card is achieved, it is better to score each subdivision.

Descriptive terms. Following are appropriate descriptive terms, with definitions, for each of the quality factors to be considered (figures in parentheses refer to the score on the 20-point score sheet)[9]:

1. Appearance. This refers to the clarity of the wine and its freedom from sediment. It is to be noted that sediment in the bottle is not the most important factor to be considered unless it is present in such an amount or form that successful decantation is impossible, or if it is due to bacterial contami-

[9] For a more complete description of these terms see pages 492–494.

nation, excessive oxidation, etc. A *brilliant* wine has no floating particles and is free of visible colloidal suspended material (2). A *clear* wine is also free of visible colloidal suspended material, but may have other defects which are not permanent, such as floating cork, filter-pad particles, etc., and it may lack the "sparkle" of a brilliant wine (1). A *dull* wine has a slight cloudiness, but any sediment that may be present is not usually due to this cloudiness (0). A *cloudy* wine has a sediment as well as cloudiness and is usually a fermenting or badly spoiled wine (0).

2. Color. This refers both to the hue and to the amount of color. However, the eye has difficulty in distinguishing hue in wines with considerable color. Appearance should not influence the score on color.

For white wines: *straw* or *colorless* wines are wines with very little color; wines with a yellow tint vary from *low* to *medium yellow* or *medium gold*. *Green* is a modifying tint. *Brown* is also a modifying tint which gives *amber*-colored wines; such wines may be marked *low, medium,* or *high amber*. It should be noted that a color which may be desirable for one type of wine is undesirable for another. Oxidized brown colors are particularly objectionable in dry white table wines.

For red wines: *pink* is the lightest shade, and *low, medium,* and *dark red* indicate a more intense color. *Violet* is a common modifying tint in young wines, particularly of certain varieties. *Brown* may also be a modifying tint in old wines. Some old red dessert wines may be *tawny*. Again, the color must be appropriate for the type of wine.

3. Odors. Odors may be divided into three types:

(a) Off-odors. *Sulfur dioxide, hydrogen sulfide, mousiness, oxidized* (or *aldehyde), corky, hot, caramel, raisin, earthy, green, rubbery, corked, woody, moldy, bacterial,* and similar terms are often used. For some wines a *baked* or *rancio* odor is undesirable, while for others it is desirable. These should be true smells and not tastes, and therefore should be recorded before tasting (if off-odors are detected, subtract 1–2).

(b) Aroma and bouquet. If there is any varietal aroma, it should be designated here with the proper variety name, such as Cabernet, Concord, Muscat, Riesling, Semillon, etc., and the degree of varietal character should be indicated (medium Riesling, low Semillon, etc.) (4). If the variety name is not known, but if the aroma is believed to be due to a specific variety—that is, it does not seem to be an off-odor or a fermentation odor—it is described as *distinct* (3). If there is no distinguishable aroma, the odor may be said to be only *vinous* (2). *Bouquet* is a term used to describe a particular odor of aged wines (1–2). Terms frequently used to describe a particular aroma or bouquet are: *rancio, foxy, flor-sherry, cooked,* etc.

(c) Volatile acidity. The volatile acid of young wines is more readily recognized and is more obejctionable than that of old wines. Gassiness is an interfering factor in the detection and estimation of the volatile acidity. Shaking will help to rid the wine of its gassiness. Both of these variables should be considered in trying to determine the actual volatile acid of the sample by tasting. Young sweet wines also show their volatile acidity readily. The following terms and point scoring should be generally applicable: *very high*, the volatile acidity is above 0.140 per cent (0); *high*, 0.101 to 0.140 per cent (0); *medium*, 0.051 to 0.100 per cent (1); and *low*, below 0.051 per cent (2). The actual score given depends on the age and type of wine.

4. Total acidity. Wines with insufficient total acid will taste *flat* and *unpalatable*; wines with too much acid will taste *acidulous, unbalanced,* or *unripe.* Even though the acidity and alcohol and body are balanced so as to give the appearance of medium acidity, an effort should be made to determine, as nearly as possible, the actual acidity by tasting. The following terms and scoring limits are suggested:

	Table wines	Dessert wines
Very high acidity, above 0.85 per cent....	0–1	0
High acidity, 0.66 to 0.85 per cent........	2	0–1
Medium acidity, 0.51 to 0.65 per cent.....	1	2
Low acidity, 0.21 to 0.50 per cent........	0	1

There must be a balance between the total acid and the alcohol and body of the wine. Thus, different types of wines may have various amounts of total acids and yet receive a full score. The most common defect in dry white table wines in California is too low a total acidity.

5. Dryness or sweetness. Dry table-wine types should normally leave no impression of sugar on the palate. The amount of sugar a wine may contain and still give the impression of being dry is affected by the total acidity, glycerine, and alcohol. A number of slightly sweet table wines are now being produced. If distinctively labeled, they should not be penalized.

Dry indicates that no impression of sweetness is made; analytically, this would usually mean less than 0.2 per cent sugar as determined by one of the ordinary copper-reduction methods. *Nearly dry* indicates that the sugar content is less than 1.0 per cent. It may or may not be unpalatable. *Sweetish* is used to describe an unpleasant or disagreeable tinge of sugar in a dry table wine, particularly in a red. Sweet wines should be judged for the actual amount of sugar present, or they may be marked *low, medium,* or *high sugar.* A tentative suggestion is made for the following limits: *low,* 0.2 to 1.0 per cent sugar (dry sherries, California dry sauterne, and *brut* champagnes); *medium,* 1.0 to 4.0 per cent sugar (California sherries and California sauterne); and *high* 4 to 15 per cent sugar (muscatel, Angelica, sweet sherry, port, Sauternes, tokay, etc.)

In scoring dry wines, an attempt must be made to determine the objectionableness of slight sugar contents. If the wine is gassy, and gives evidence of fermenting and giving a sediment, even the slightest sugar content may be objectionable (0). For sparkling wine the terms *brut* or *nature* are used for dry, *sec* for low sugar, and *demi-sec* for medium sugar.

6. Body. The comparison of the extract content is not easy when wines of varying color, sugar, and alcohol are being tasted. For dry wines the following range is suggested: *low,* extract below 2.0 per cent (light dry white wine types fall in this class); *medium,* extract from 2.0 to 3.0 per cent; and *high,* extract above 3.0 per cent (many red table wines are high in body). Other terms used to describe the body are *watered, thin,* or *heavy.* All dessert wines score (1) as the body is masked by the sugar.

7. Flavors. When a descriptive term can be found it should be used. The following terms are among those commonly used: *fruity, bitter, gassy, metallic, alcoholic, dry* (due to overaging). Strictly speaking, probably the only true taste listed is *bitter,* but in the organoleptic examination of

wines these so-called flavors only become obvious when the wine is placed in the mouth. Terms such as *fermenting, young, becoming mature, mature* (*bottle-ripe* or *ready to drink*), *aged,* and *very old* (*passé*) may be conveniently used to describe the maturation of a wine. Most wines score (2) unless they have very marked defects or merits.

8. Astringency. The chief factors interfering with the taste of tannin are acidity and gassiness, but alcohol, sweetness, and body may also interfere. Usually there is a direct relationship between the amount of red color and the tannin determination. For red wines the tannin varies from about 0.1 to 0.3 per cent. Most white wines have a tannin content of less than 0.05 per cent. If above these figures, the tannin is high for a white wine and the wine is usually too astringent (0). The impression of smoothness is considerably affected by the tannin, glycerine, and sugar contents of a sample.

An approximate classification can be made as follows: *smooth,* the wine has no trace of harshness; usually used for aged, white table or sweet wines (2); *slightly rough,* the pink or red wine has a trace of roughness on the palate, due to lack of aging or to tannin in the wine (1–2); *medium rough,* the red wine is ordinarily not aged (0–1); *rough,* new red wines often are harsh and rough to the palate (0); and *very rough,* a particularly rough red wine, due possible to the variety (0). Other terms related to these are *soft, rounded, mellow, harsh, hard.*

9. General quality. Although all the factors influencing the total score should have been taken care of under the headings discussed above, it is sometimes desirable to adjust the total score to the quality groups listed below by ascribing to the wine a general quality factor. Wines of poor overall quality score 0, while wines giving a very favorable over-all impression score 2. The justification for this is that the quality may involve the interrelationship between components as well as each individual one.

Scoring. After evaluating the above-mentioned factors for a particular type of wine, the total score should indicate the quality of the wine. The wines are always scored as they are at the time of tasting. Young wines may give promise of becoming superior or even fine wines, but they cannot usually be so ranked because they lack age and proper smoothness. The following groupings are suggested: (a) *good* (16–20 points), wines of fine quality, which should be balanced, have no noticeable or pronounced defect, and, if bottled, be free of "young" character; (b) *standard* (11–15 points), the wines of commerce in California (including ordinary wines in bottle), which are not deficient in any important characteristic but lack finesse, age, character, or the balance required for bottling or fine quality; (c) *common* (6–10 points), wines lacking some necessary requirement or suffering from some malady (wines with an off-taste or high volatile acidity fall here); (d) *spoiled* (1–5 points), wines so spoiled that they must be discarded.

Thurstone (1952) has pointed out the advantage of hedonic scales of evaluation. These are based on the pleasure or displeasure of the reaction of the taster to the wine. For this kind of evaluation the taster need have no preconceived notions of quality. For consumer tests this has a great advantage, since the hedonic reactions may be converted to numerical values that can be interpreted by the usual statistical procedures. See Guilford (1954) for information on the statistical analyses of such data.

TABLE 7

COMPARISON OF QUALITY RATINGS, SCORING SYSTEMS, AND DECRIPTIVE CLASSIFICATIONS

Quality	Hedonic score	10-pt. scale	20-pt. scale	Description	Hedonic rating
Excellent	7	10	18–20	Bottle bouquet, perfect appearance and taste.	One of the best wines I have tasted; like extremely.
Good	6	9	15–17	Aged bouquet, nearly perfect appearance or taste.	Very pleasing; like very much.
Fair	5	7–8	12–14	Distinct bouquet, nearly perfect appearance or taste.	Slightly pleasing; like mildly.
Ordinary	4	5–6	9–11	Distinct aroma, nearly perfect appearance or taste.	Neither pleasant nor unpleasant; neither like nor dislike.
Poor	3	4–3	6–8	Less distinct aroma, defective appearance or taste.	Slightly unpleasant; dislike mildly.
Bad	2	2	3–5	Slight defect in odor, very defective appearance or taste.	Distinctly unpleasant; dislike very much.
Very bad	1	1	1–2	Defective in odor and imperfect apperance and taste.	One of the most unpleasant wines; dislike extremely.

TABLE 8

A HEDONIC SCORE CARD FOR WINES

Type..................................

Appearance	Very desirable	Desirable	Acceptable	Not acceptable	Very undesirable
Color	Very desirable	Desirable	Acceptable	Not acceptable	Very undesirable
Aroma	Very pleasing	Pleasing	Neither pleasing nor unpleasing		
Odors	Very pleasing	Pleasing	Neither pleasing nor unpleasing	Slightly unpleasant	Very unpleasant
Sugar	Very pleasing	Pleasing	Neither pleasing nor unpleasing	Slightly unpleasant	Very unpleasant
Acidity	Very pleasing	Pleasing	Neither pleasing nor unpleasing	Slightly unpleasant	Very unpleasant
Smoothness or bitterness	Very pleasing	Pleasing	Neither pleasing nor unpleasing	Slightly unpleasant	Very unpleasant
General impression	Very pleasing	Pleasing	Neither pleasing nor unpleasing	Slightly unpleasant	Very unpleasant

The panel should have a clear concept of the meaning of the quality rating of each step if comparable results between different judges are to be obtained. Definitions to be used with a 10- or 20-point scale are suggested in table 7. The hedonic scale is highly subjective and whether it can be compared with a quality scale is open to question.

One of the advantages of hedonic scales as compared with numerical scales is that the hedonic terms constitute a definition of each step of the scale. Hedonic scales are particularly useful for untrained judges or for judges who are not accustomed to using any particular numerical scale. For a description of various scales see Jones *et al.* (1955).

Koch (1955) devised the following score card for fruit wines:

Clarity	
Clear	2
Slightly cloudy	1
Cloudy	0
Color	
Normal	2
Too much color	1
Off-color or lack of color	0
Smell	
Correct	3
Foreign	1
Off	0
Taste	
Correct	9–10
Harmonious	7–8
Good	5–6
Inharmonious	3–4
Foreign	1–2
Off	0

Koch also gives points for a chemical scale: 3 points for a volatile acidity below 0.06 gm per 100 ml and a sulfur dioxide of less than 15 mg per liter free and 100 mg per liter total; 2 points for less than 0.08 gm per 100 ml volatile acidity and less than 30 mg per liter of free and less than 150 mg per liter total sulfur dioxide; 1 point, a maximum of 0.10 gm per 100 ml volatile acidity and not more than 50 mg per liter free and less than 200 mg per liter total sulfur dioxide; and 0 points for more than 0.10 gm per 100 ml volatile acidity and more than 50 mg per liter free and 200 mg per liter total sulfide dioxide. Therefore, as many as 17 points may be given for the sensory examination and as many as 3 for the chemical.

There is no evidence that the suggested scores are actually related to the organoleptic quality, but his curve for scores of 0 to 20 is a fairly straight line. The maximum score of 20 is assigned thus: 1 to 8 points, deficient, 0 per cent; 9 to 11 points, 0 to 24 per cent of maximum; 12 to 14 points, average, 25 to 50 per cent; 15 to 17 points, 50 to 75 per cent; and 18 to 20 points, 75 to 100 per cent. Obviously this scale has only about 13 steps because all scores below 9 are given a quality score of 0 per cent. This is not as irrational as it seems, since there may be a range of quality below that for commercial acceptability.

TABLE 9

DISTRIBUTION OF F AND t

5 per cent level

n_2	n_1										t
	1	2	3	4	5	6	8	12	24	∞	
1	161.40	199.50	215.70	224.60	230.20	234.00	238.90	243.90	249.00	254.30	12.71
2	18.51	19.00	19.16	19.25	19.30	19.33	19.37	19.41	19.45	19.50	4.30
3	10.13	9.55	9.28	9.12	9.01	8.94	8.84	8.74	8.64	8.52	3.18
4	7.71	6.94	6.59	6.39	6.26	6.16	6.04	5.91	5.77	5.63	2.78
5	6.61	5.79	5.41	5.19	5.05	4.95	4.82	4.68	4.53	4.36	2.57
6	5.99	5.14	4.76	4.53	4.39	4.28	4.15	4.00	3.84	3.67	2.45
7	5.59	4.74	4.35	4.12	3.97	3.87	3.73	3.57	3.41	3.23	2.36
8	5.32	4.46	4.07	3.84	3.69	3.58	3.44	3.28	3.12	2.93	2.31
9	5.12	4.26	3.86	3.63	3.48	3.37	3.23	3.07	2.90	2.71	2.26
10	4.96	4.10	3.71	3.48	3.33	3.22	3.07	2.91	2.74	2.54	2.23
11	4.84	3.98	3.59	3.36	3.20	3.09	2.95	2.79	2.61	2.40	2.20
12	4.75	3.88	3.49	3.26	3.11	3.00	2.85	2.69	2.50	2.30	2.18
13	4.67	3.80	3.41	3.18	3.02	2.92	2.77	2.60	2.42	2.21	2.16
14	4.60	3.74	3.34	3.11	2.96	2.85	2.70	2.53	2.35	2.13	2.14
15	4.54	3.68	3.29	3.06	2.90	2.79	2.64	2.48	2.29	2.07	2.13
16	4.49	3.63	3.24	3.01	2.85	2.74	2.59	2.42	2.24	2.01	2.12
17	4.45	3.59	3.20	2.96	2.81	2.70	2.55	2.38	2.19	1.96	2.11
18	4.41	3.55	3.16	2.93	2.77	2.66	2.51	2.34	2.15	1.92	2.10
19	4.38	3.52	3.13	2.90	2.74	2.63	2.48	2.31	2.11	1.88	2.09
20	4.35	3.49	3.10	2.87	2.71	2.60	2.45	2.28	2.08	1.84	2.09
30	4.17	3.32	2.92	2.69	2.53	2.42	2.27	2.09	1.89	1.62	2.04
40	4.08	3.23	2.84	2.61	2.45	2.34	2.18	2.00	1.79	1.51	2.02
60	4.00	3.15	2.76	2.52	2.37	2.25	2.10	1.92	1.70	1.39	2.00
120	3.92	3.07	2.68	2.45	2.29	2.17	2.02	1.83	1.61	1.25	1.98
∞	3.84	2.99	2.60	2.37	2.21	2.10	1.94	1.75	1.52	1.00	1.96

Score cards often omit any general quality factor. They therefore neglect what many experienced tasters find, namely, that the balance or harmonious complex of smell and taste is itself an important part of the quality impression. As a matter of fact, competent critics may have a very elementary conception of the physiology of taste and smell or of the nature of the four basic tastes. They have acquired through practice a recognition of the desirable harmonious interrelationships of the various tastes.

It is possible that with trained tasters, a hedonic score card may yield useful information if the impressions for each characteristic are classified. This was suggested to us by the score card for meat developed by Satorius and Child (1938). A tentative five-step hedonic score card for wine, with hedonic scales for each character, is given in table 8.

The taster must first be told what the type of wine is, or he should be expected to determine the type before completing the score card.

Analyses of Results: One of the chief disadvantages of score cards or hedonic systems is that it is more troublesome to determine the significance of differences in scores. Furthermore, to use the usual statistical procedures the scores should have a normal distribution. Amerine and Roessler (1952) showed that for 822 wines scored at Davis the distribution was only moder-

TABLE 10

DISTRIBUTION OF F AND t

1 per cent level

n_2	n_1: 1	2	3	4	5	6	8	12	24	∞	t
1	4052	4999	5403	5625	5764	5859	5982	6106	6234	6366	63.66
2	98.50	99.00	99.17	99.25	99.30	99.33	99.37	99.42	99.46	99.50	9.92
3	34.12	30.82	29.46	28.71	28.24	27.91	27.49	27.05	26.60	26.12	5.84
4	21.20	18.00	16.69	15.98	15.52	15.21	14.80	14.37	13.93	13.46	4.60
5	16.26	13.27	12.06	11.39	10.97	10.67	10.29	9.89	9.47	9.02	4.03
6	13.74	10.92	9.78	9.15	8.75	8.47	8.10	7.72	7.31	6.88	3.71
7	12.25	9.55	8.45	7.85	7.46	7.19	6.84	6.47	6.07	5.65	3.50
8	11.26	8.65	7.59	7.01	6.63	6.37	6.03	5.67	5.28	4.86	3.36
9	10.56	8.02	6.99	6.42	6.06	5.80	5.47	5.11	4.73	4.31	3.25
10	10.04	7.56	6.55	5.99	5.64	5.39	5.06	4.71	4.33	3.91	3.17
11	9.65	7.20	6.22	5.67	5.32	5.07	4.74	4.40	4.02	3.60	3.11
12	9.33	6.93	5.95	5.41	5.06	4.82	4.50	4.16	3.78	3.36	3.06
13	9.07	6.70	5.74	5.20	4.86	4.62	4.30	3.96	3.59	3.16	3.01
14	8.86	6.51	5.56	5.03	4.69	4.46	4.14	3.80	3.43	3.00	2.98
15	8.68	6.36	5.42	4.89	4.56	4.32	4.00	3.67	3.29	2.87	2.95
16	8.53	6.23	5.29	4.77	4.44	4.20	3.89	3.55	3.18	2.75	2.92
17	8.40	6.11	5.18	4.67	4.34	4.10	3.79	3.45	3.08	2.65	2.90
18	8.28	6.01	5.09	4.58	4.25	4.01	3.71	3.37	3.00	2.57	2.88
19	8.18	5.93	5.01	4.50	4.17	3.94	3.63	3.30	2.92	2.49	2.86
20	8.10	5.85	4.94	4.43	4.10	3.87	3.56	3.23	2.86	2.42	2.84
30	7.56	5.39	4.51	4.02	3.70	3.47	3.17	2.84	2.47	2.01	2.75
40	7.31	5.18	4.31	3.83	3.51	3.29	2.99	2.66	2.29	1.80	2.70
60	7.08	4.98	4.13	3.65	3.34	3.12	2.82	2.50	2.12	1.60	2.66
120	6.85	4.79	3.95	3.48	3.17	2.96	2.66	2.34	1.95	1.38	2.62
∞	6.64	4.60	3.78	3.32	3.02	2.80	2.51	2.18	1.79	1.00	2.58

ately asymmetrical, and they attributed this to the unfinished nature of the wines tasted. For the discussion that follows we will assume that the wines have been scored on a 20-point scale and that the distributions are normal. The steps for determining whether two scores are significantly different from each other depend upon the tasting design.

Paired scores: Statistical tests for differences between wines are based on the assumption that no differences exist in the populations from which the samples are drawn. This assumption is known as the null hypothesis and applies not only to population mean scores but also to population standard deviations.

If one taster in comparing two wines at several tastings scores a sample of each, these scores are paired; or if members of a panel of tasters each taste and score a sample of each of two wines, a set of paired scores results. In such a situation the t-distribution for paired samples furnishes a test of significance of the mean difference in scores between the two wines. In the case of two wines with n paired scores, X and Y, the differences $D = X - Y$ are analyzed and the difference $\overline{D}$ between mean scores $\overline{X}$ and $\overline{Y}$, $\overline{D} = \overline{X} - \overline{Y}$, may be tested by calculating $t = \overline{D}/s_{\bar{d}}$ (with $n - 1$ degrees of freedom) where $s_{\bar{d}}$, the

best estimate of the standard error of the population of differences between the means of scores, is calculated from the expression

$$s_{\bar{d}} = \frac{s}{\sqrt{n}}$$

where

$$s = \sqrt{\frac{\Sigma(D - \bar{D})^2}{n - 1}} = \sqrt{\frac{\Sigma D^2 - \frac{(\Sigma D)^2}{n}}{n - 1}} .$$

The significance of the result is determined by comparing the calculated value of t with the tabular values (tables 9 and 10 for the appropriate number of degrees of freedom).

As an example, suppose that seven tasters scored each of two wines. The results were as follows:

Tasters	Scores X	Y	D	D^2
A	15	14	1	1
B	12	14	−2	4
C	14	15	−1	1
D	17	14	3	9
E	11	11	0	0
F	16	14	2	4
G	15	13	2	4
Totals	100	95	5	23
Means	14.28	13.57	0.71	

$$s = \sqrt{\frac{23 - \frac{(5)^2}{7}}{6}} = \sqrt{3.24}$$

$$s_{\bar{d}} = \sqrt{\frac{3.24}{7}} = \sqrt{0.46} = 0.68$$

$$t = \frac{0.71}{0.68} = 1.04 \text{ (with 6 degrees of freedom)}$$

At the 5 per cent level of significance with 6 degrees of freedom, the tabular value of t is 2.45. Since the calculated value of t is less than that expected at the 5 per cent level of significance, the difference between the mean scores of the wines is not significant.

Instead of comparing calculated and tabular values of t, significance between mean scores is sometimes determined by comparing the actual differences between means with the difference required for significance as calculated from $\bar{D}' = t' s_{\bar{d}}$ where t', as before, is the tabular value of t, for $n - 1$ degrees of freedom, at the required level of significance. In the example, at the 5 per cent level of significance, $\bar{D}' = 2.45\ (0.68) = 1.7$. Since the difference between mean scores is only 0.71, it is not significant.

Scores Not Paired: If the scores are not paired—that is, if to any score of one wine no particular score of the other corresponds—then the t-distribution still furnishes the appropriate test for significance. In this case for n_1 X-scores and n_2 Y-scores (n_1 may equal n_2),

$$t = \frac{\bar{X} - \bar{Y}}{s_{\bar{x}-\bar{y}}} \quad \text{with } (n_1 - 1) + (n_2 - 1) \text{ degrees of freedom}$$

where

$$s_{\bar{x}-\bar{y}} = \sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}} \quad \text{and} \quad s^2 = \frac{\Sigma X^2 - \frac{(\Sigma X)^2}{n_1} + \Sigma Y^2 - \frac{(\Sigma Y)^2}{n_2}}{(n_1 - 1) + (n_2 - 1)}.$$

As before, if the calculated value of t exceeds the tabular value for $(n_1 - 1) + (n_2 - 1)$ degrees of freedom, the mean scores are significantly different. In this case the actual differences between means may be compared with the difference required for significance as calculated from $\bar{D}' = t's_{\bar{x}-\bar{y}}$ where t' is again the tabular value of t at the appropriate level of significance.

Several Wines Scored by Each of Several Tasters: In the usual tasting design in which a panel of k tasters scores each of n wines, the analysis of variance procedure yields the most information. In such an analysis the total sum of squares based upon the variation of all scores is broken down into three parts: a sum of squares based upon variation between the wines; a sum of squares based upon variation between the tasters; and a sum of squares which is not the result of variation between wines or tasters. This last sum of squares is known as the remainder or error sum of squares. It is a measure of the unexplained variation in the experiment. The number of degrees of freedom is broken down in the same way. This is known as a randomized block design whose pattern for analysis is indicated below:

Tasters	Wines					
	1	2	3	. . .	n	Totals
1	X_1	Y_1	Z_1	. . .	V_1	T_1
2	X_2	Y_2	Z_2	. . .	V_2	T_2
3	X_3	Y_3	Z_3	. . .	V_3	T_3
.	.	.	.			
.	.	.	.			
.	.	.	.			
k	X_k	Y_k	Z_k	. . .	V_k	T_k
Totals	W_1	W_2	W_3	. . .	W_n	$\Sigma T = \Sigma W$

Computations:

(a) Correction term (c.t.) $= \dfrac{(\Sigma T)^2}{nk}$

(b) Total sum of squares $= \Sigma X^2 + \Sigma Y^2 + \Sigma Z^2 + \cdots + \Sigma V^2 - \text{c.t.}$

(c) Sum of squares for wines $= \dfrac{\Sigma W^2}{k} - \text{c.t.}$

TABLE 11

SIGNIFICANT STUDENTIZED RANGES
MULTIPLE RANGE TEST
5 per cent level

d.f.	p 2	3	4	5	6	7	8	9	10	20	100
1	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0
2	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09
3	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50	4.50
4	3.93	4.01	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02	4.02
5	3.64	3.74	3.79	3.83	3.83	3.83	3.83	3.83	3.83	3.83	3.83
6	3.46	3.58	3.64	3.68	3.68	3.68	3.68	3.68	3.68	3.68	3.68
7	3.35	3.47	3.54	3.58	3.60	3.61	3.61	3.61	3.61	3.61	3.61
8	3.26	3.39	3.47	3.52	3.55	3.56	3.56	3.56	3.56	3.56	3.56
9	3.20	3.34	3.41	3.47	3.50	3.52	3.52	3.52	3.52	3.52	3.52
10	3.15	3.30	3.37	3.43	3.46	3.47	3.47	3.47	3.47	3.48	3.48
11	3.11	3.27	3.35	3.39	3.43	3.44	3.45	3.46	3.46	3.48	3.48
12	3.08	3.23	3.33	3.36	3.40	3.42	3.44	3.44	3.46	3.48	3.48
13	3.06	3.21	3.30	3.35	3.38	3.41	3.42	3.44	3.45	3.47	3.47
14	3.03	3.18	3.27	3.33	3.37	3.39	3.41	3.42	3.44	3.47	3.47
15	3.01	3.16	3.25	3.31	3.36	3.38	3.40	3.42	3.43	3.47	3.47
16	3.00	3.15	3.23	3.30	3.34	3.37	3.39	3.41	3.43	3.47	3.47
17	2.98	3.13	3.22	3.28	3.33	3.36	3.38	3.40	3.42	3.47	3.47
18	2.97	3.12	3.21	3.27	3.32	3.35	3.37	3.39	3.41	3.47	3.47
19	2.96	3.11	3.19	3.26	3.31	3.35	3.37	3.39	3.41	3.47	3.47
20	2.95	3.10	3.18	3.25	3.30	3.34	3.36	3.38	3.40	3.47	3.47
30	2.89	3.04	3.12	3.20	3.25	3.29	3.32	3.35	3.37	3.47	3.47
40	2.86	3.01	3.10	3.17	3.22	3.27	3.30	3.33	3.35	3.47	3.47
60	2.83	2.98	3.08	3.14	3.20	3.24	3.28	3.31	3.33	3.47	3.48
100	2.80	2.95	3.05	3.12	3.18	3.22	3.26	3.29	3.32	3.47	3.53
∞	2.77	2.92	3.02	3.09	3.15	3.19	3.23	3.26	3.29	3.47	3.67

(d) Sum of squares for tasters $= \dfrac{\Sigma T^2}{n} -$ c.t.

(e) Remainder (error) sum of squares = (b) − (c) − (d).

Degrees of freedom:

Total	$nk - 1$
Wines	$n - 1$
Tasters	$k - 1$
Remainder	$(nk - 1) - (n - 1) - (k - 1) = (n - 1)(k - 1)$

From these sums of squares and the corresponding numbers of degrees of freedom, three independent estimates of the population variance are computed. On the assumption that the groups making up the total set of measurements are random samples from a homogeneous population, the three estimates of the population variance may be expected to differ only within

TABLE 12

SIGNIFICANT STUDENTIZED RANGES
MULTIPLE RANGE TEST
1 per cent level

d.f.	p										
	2	3	4	5	6	7	8	9	10	20	100
1	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0
2	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0
3	8.26	8.5	8.6	8.7	8.8	8.9	8.9	9.0	9.0	9.3	9.3
4	6.51	6.8	6.9	7.0	7.1	7.1	7.2	7.2	7.3	7.5	7.5
5	5.70	5.96	6.11	6.18	6.26	6.33	6.40	6.44	6.5	6.8	6.8
6	5.24	5.51	5.65	5.73	5.81	5.88	5.95	6.00	6.0	6.3	6.3
7	4.95	5.22	5.37	5.45	5.53	5.61	5.69	5.73	5.8	6.0	6.0
8	4.74	5.00	5.14	5.23	5.32	5.40	5.47	5.51	5.5	5.8	5.8
9	4.60	4.86	4.99	5.08	5.17	5.25	5.32	5.36	5.4	5.7	5.7
10	4.48	4.73	4.88	4.96	5.06	5.13	5.20	5.24	5.28	5.55	5.55
11	4.39	4.63	4.77	4.86	4.94	5.01	5.06	5.12	5.15	5.39	5.39
12	4.32	4.55	4.68	4.76	4.84	4.92	4.96	5.02	5.07	5.26	5.26
13	4.26	4.48	4.62	4.69	4.74	4.84	4.88	4.94	4.98	5.15	5.15
14	4.21	4.42	4.55	4.63	4.70	4.78	4.83	4.87	4.91	5.07	5.07
15	4.17	4.37	4.50	4.58	4.64	4.72	4.77	4.81	4.84	5.00	5.00
16	4.13	4.34	4.45	4.54	4.60	4.67	4.72	4.76	4.79	4.94	4.94
17	4.10	4.30	4.41	4.50	4.56	4.63	4.68	4.72	4.75	4.89	4.89
18	4.07	4.27	4.38	4.46	4.53	4.59	4.64	4.68	4.71	4.85	4.85
19	4.05	4.24	4.35	4.43	4.50	4.56	4.61	4.64	4.67	4.82	4.82
20	4.02	4.22	4.33	4.40	4.47	4.53	4.58	4.61	4.65	4.79	4.79
30	3.89	4.06	4.16	4.22	4.32	4.36	4.41	4.45	4.48	4.65	4.71
40	3.82	3.99	4.10	4.17	4.24	4.30	4.34	4.37	4.41	4.59	4.69
60	3.76	3.92	4.03	4.12	4.17	4.23	4.27	4.31	4.34	4.53	4.66
100	3.71	3.86	3.98	4.06	4.11	4.17	4.21	4.25	4.29	4.48	4.65
∞	3.64	3.80	3.90	3.98	4.04	4.09	4.14	4.17	4.20	4.41	4.68

the limits of chance fluctuations. This is the null hypothesis and is tested by comparing the variances between wines and between tasters with the error variance. In each case the comparison consists of calculating a variance ratio

$$F = \frac{\text{larger variance}}{\text{smaller variance}}.$$

Values of F to be expected at the 5 per cent and 1 per cent levels of significance are given in tables 9 and 10. If the calculated value of F exceeds the tabular value for the level of significance agreed upon, the null hypothesis—namely, that there is no difference among the populations from which the samples have been drawn—is considered untenable. If the null hypothesis is rejected, the populations from which the samples have been drawn may differ in terms of either means or variances or both. If the variances are approximately the same, it is the means that differ. In experimental data it is usually the means and not the variances that disagree.

The F-distribution provides an over-all test of significance among the different means. If the F-test indicates differences, then specific comparisons may be made by using the least significant difference based on the t-distribu-

tion or by applying one of the newer tests. An appropriate procedure is to use the multiple-range test as outlined by Duncan (1955) to set up a series of shortest significant ranges with which to compare differences between mean scores. The shortest significant range R_p for comparing the largest and the smallest of p mean scores is given by $R_p = Qs_{\bar{x}}$ with $(n-1)(k-1)$ degrees of freedom where the appropriate values of Q can be obtained from tables 11 and 12. Q is the significant studentized ranges at the 5 and 1 per cent levels, and $s_{\bar{x}}$ is the standard error of any mean score, that is, $s_{\bar{x}} = \sqrt{\text{error variance}/k}$.

To illustrate, the test will be applied to scores (fictitious) assigned by 10 tasters to three different wines. Differences between the mean scores will be tested for significance.

Tasters	Wines			
	X	Y	Z	Totals
A	13	14	13	40
B	14	15	12	41
C	14	15	13	42
D	12	14	12	38
E	14	14	12	40
F	15	13	13	41
G	13	15	12	40
H	13	12	12	37
I	14	14	12	40
J	15	14	13	42
Totals	137	140	124	401
Means	13.7	14.0	12.4	

Computations:

(a) Correction term (c.t.) $= \dfrac{(401)^2}{30} = 5360.03$

(b) Total sum of squares $= (13)^2 + (14)^2 + \cdots + (13)^2 - \text{c.t.}$
$= 5393.00 - 5360.03 = 32.97$

(c) Sum of squares for wines
$= \dfrac{(137)^2 + (140)^2 + (124)^2}{10} - \text{c.t.} = 5374.50 - 5360.03 = 14.47$

(d) Sum of squares for tasters
$= \dfrac{(40)^2 + (41)^2 + \cdots + (42)^2}{3} - \text{c.t.} = 5367.67 - 5360.03 = 7.64$

(e) Remainder (error) sum of squares = 32.97 − 22.11 = 10.86.

Analysis of variance

Source of variation	Degrees of freedom	Sum of squares	Mean square (variance)	*F* value		
				Calculated	Tabular	
					5%	1%
Total	29	32.97	.....	....	....	...
Wines	2	14.47	7.24	12.0	3.55	6.01
Tasters	9	7.64	0.849	1.41	2.46	...
Error	18	10.86	0.603	....	....	....

Since the calculated value of F (12.0) is larger than the tabular values (3.55, 6.01), the analysis indicates, at both the 5 per cent and 1 per cent levels, significant differences between mean scores of wines. No significant differences between tasters are indicated.

To determine which mean scores are significantly different at the 5 per cent level, the values of Q, for 18 degrees of freedom, to be used in comparing the largest and smallest values of sets of two ($p = 2$) and three ($p = 3$) means are selected. They are 2.97 and 3.12 respectively. For the 1 per cent level of significance, table 12 is employed. The studentized ranges Q are each multiplied by the standard error $\sqrt{0.603/10} = 0.246$ to form the shortest significant ranges R_p. Adjacent means having a difference greater than R_2 are considered significantly different. Differences between the smallest and largest of three means arranged in order of size are significant if they exceed R_3, or in general R_p represents the dividing line between significance and nonsignificance for the smallest and largest of p means arranged in order of size. The results are summarized in the following table.

Multiple Range Test

	Shortest Significant Ranges			
	5% level		1% level	
	$p = 2$	$p = 3$	$p = 2$	$p = 3$
Q	2.97	3.12	4.07	4.27
R_p	0.73	0.77	1.00	1.05

Comparisons

Wines	Z	X	Y
Mean scores	12.4	$\underline{13.7}$	$\underline{14.0}$

Note: Any two means *underscored* by the same line are *not significantly different*.
Any two means *not underscored* by the same line are *significantly different*.

Differences between mean scores are compared with R_p as follows.

	5% level	1% level
$Y - Z = 1.6 >$	.77 sig.	$1.6 > 1.05$ sig.
$Y - X = 0.3 <$	.73 not sig.	$0.3 < 1.0$ not sig.
$X - Z = 1.3 >$	.73 sig.	$1.3 > 1.0$ sig.

Therefore, at both the 5 per cent and 1 per cent levels of significance the mean scores of wines X and Y are not significantly different, but both are significantly better than that of wine Z. Comparisons are indicated by underscoring as shown above.

If there are available merely sets of scores for each of n wines, the above procedure must be modified. Suppose there are k_1 scores for wine X, k_2 for wine Y, k_3 for wine Z, . . . , k_n for wine V. The k's may be the same or different.

Then in the general analysis of variance pattern the correction term (a) and the total sum of squares (b) are calculated as before, but a modification must be made in the sum of squares for wines (c) if the k's are not the same, and the sum of squares for tasters (d) is eliminated. In this case

(c) Sum of squares for wines $= \frac{W_1^2}{k_1} + \frac{W_2^2}{k_2} + \frac{W_3^2}{k_3} + \cdots + \frac{W_n^2}{k_n} - \text{c.t.}$

(d) Sum of squares for tasters $= 0$

(e) Remainder (error) sum of squares = (b) − (c) with $\Sigma k - n$ degrees of freedom.

If the mean scores are based on different numbers of individual scores—that is, if the k's are different—an effective number of replications applicable to all mean scores must be determined before calculating $s_{\bar{x}}$. Such an effective number of replications may be determined from the formula

$$k_o = \frac{1}{n-1}\left(\Sigma k - \frac{\Sigma k^2}{\Sigma k}\right)$$

and then

$$s_{\bar{x}} = \sqrt{\frac{\text{error variance}}{k_o}}\,.$$

For example, consider the scores (fictitious) of four wines as follows:

Wine Scores

	X	Y	Z	V	
	16	17	15	14	
	12	14	11	12	
	11	12	13	17	
	13	17	12	15	
	11	15	13	14	
	15	14	15	13	
	14	17	13	15	
	15	16	10		
	11	15	16		
	13	17			
	14				
	13				
Totals......	158	154	118	100	530
Means......	13.2	15.4	13.1	14.3	

Correction term (c.t.) $= \frac{(530)^2}{38} = 7392.11$

Total sum of squares

$= (16)^2 + (12)^2 + \cdots + (15)^2 - \text{c.t.} = 7532 - 7392.11 = 139.89$

Sum of squares for wines

$= \frac{(158)^2}{12} + \frac{(154)^2}{10} + \frac{(118)^2}{9} + \frac{(100)^2}{7} - \text{c.t.} = 7427.62 - 7392.11 = 35.51$

Remainder (error) sum of squares $= 139.89 - 35.51 = 104.38$

Analysis of variance

Source of variation	Degrees of freedom	Sum of squares	Mean square (variance)	*F* value	
				Calculated	Tabular 5%
Total	37	139.89	....	....	....
Wines	3	35.51	11.84	3.86	2.88
Error	34	104.38	3.07	....	....

Since the calculated value of *F* (3.86) exceeds the value 2.88 at the 5 per cent level, significant differences between mean scores of wines are indicated. As before, the multiple-range test will be employed to find specific differences.

Therefore

$$k_o = \frac{1}{n-1}\left(\Sigma k - \frac{\Sigma k^2}{\Sigma k}\right) = \frac{1}{4-1}\left(38 - \frac{374}{38}\right) = \frac{1}{3}(38 - 9.84) = 9.39$$

and

$$s_{\bar{x}} = \sqrt{\frac{3.07}{9.39}} = \sqrt{.327} = 0.572$$

Multiple Range Test

Shortest Significant Range 5% level

	$p = 2$	$p = 3$	$p = 4$
Q	2.88	3.03	3.11
R_p	1.65	1.73	1.78

Wines	Z	X	V	Y
	Comparisons			
Mean scores	13.1	13.2	14.3	15.4

Note: Any two mean scores *underscored* by the same line are *not significantly different*. Any two mean scores *not underscored* by the same line are *significantly different.*

Therefore, the score of wine Y is significantly better than those of wines X and Z.

Exactly the same procedure would be employed for a 10-point scale. If a seven-step hedonic scale were used, the hedonic ratings would be converted to points and analyzed in the same manner. Harries (1956) has shown that there is a positional bias, which differs between tasters, when samples are scored 0 to 7 and numbered 1 to 6. The tendency is to treat the end samples differently and to score samples higher or lower according to their position. This is different from the well-known phenomenon of scoring a sample low if it follows a better one, or high if it follows a poorer one. Harries suggested treating the code sample as a factor and including its effect in the main table of variance and disregarding it in the tests of significance of the main treatment effects. Filipello (1956*b*) reports positional bias in consumer-acceptance tests as well.

Ranking Methods. Sometimes judges rank wines in order of merit instead of assigning to each wine a numerical measure of intrinsic worth. Ranking obviously does not supply as much information as scoring, since it gives no indication of the judge's opinion of the degree by which two wines differ.

In the case of n wines which have been ranked by two judges, Spearman's Rank Correlation Coefficient, which is defined as

$$R = 1 - \frac{6\Sigma d^2}{n^3 - n},$$

where Σd^2 is the sum of squares of the differences in rank, measures the agreement between the rankings assigned by the two judges. R may vary from a value of -1, indicating a complete reversal of ranking between the judges, to $+1$, representing perfect agreement. $R = 0$ indicates the ranks to be totally unrelated. How may one be sure that the value of R did not arise by chance? In other words, how do we test whether the calculated value of R is significantly different from zero? Little reliability can be placed in a value of R determined from less than 10 wines. This fact, too, places a limitation on the usefulness of the ranking procedure. For $n > 10$, the significance of a calculated R may be determined by use of the t-test, i.e.,

$$t = R\sqrt{\frac{n-2}{1-R^2}} \quad \text{with } (n-2) \text{ d.f.}$$

Significance is determined by comparing the calculated value of t with the tabular values from tables 9 and 10.

For example, suppose two judges rank 10 wines as follows:

	Wines										
	1	2	3	4	5	6	7	8	9	10	
Ranked by A	2	1	8	7	10	3	9	4	6	5	
Ranked by B	3	2	5	9	10	1	8	4	6	7	
Diff. in ranks = d	1	1	−3	2	0	−2	−1	0	0	2	
d^2	1	1	9	4	0	4	1	0	0	4	$\Sigma d^2 = 24$

Then

$$R = 1 - \frac{6(24)}{1000 - 10} = 0.854$$

and

$$t = 0.854\sqrt{\frac{8}{1 - 0.729}} = 0.854\ (5.43)$$

$$= 4.64 \text{ (with 8 d.f.)}$$

The tabular value of t for 8 degrees of freedom at the 1 per cent level of significance is 3.36. Therefore, the agreement in ranking by the two judges is highly significant.

If the correct ranking is known according to some characteristic, this method may be employed to test whether an individual is really discriminating and would make an acceptable judge.

In testing for significant agreement between the rankings assigned n wines by k judges instead of two judges, the data are arranged in an analysis of variance pattern, and a statistic W, the coefficient of concordance, is calculated. The sampling distribution of W under the null hypothesis has been investigated by Kendall (1948), who has shown that W may be tested for significance by use of the F-distribution. For small values of k continuity corrections are appropriate. The value of the coefficient of concordance, corrected for continuity, is given by

$$W_c = \frac{\text{s.s. between wines} - 1/k}{\text{total s.s.} + 2/k}$$

and

$$F = \frac{(k-1)W_c}{1 - W_c}$$

with the appropriate degrees of freedom estimated from the expression

$$\text{d.f. for numerator} \doteq (n-1) - \frac{2}{k}$$

$$\text{d.f. for denominator} = (k-1)\left[(n-1) - \frac{2}{k}\right]$$

As an example, consider five wines ($n = 5$) ranked by 4 judges ($k = 4$). As shown in the table,

	Rank of wines					
Judges	X	Y	Z	U	V	
A..........	4	2	3	5	1	
B..........	3	4	1	5	2	
C..........	1	2	4	5	3	
D..........	5	3	2	4	1	
Totals......	13	11	10	19	7	60

Then

$$\text{c.t.} = \frac{(60)^2}{20} = 180$$

$$\begin{aligned}\text{s.s. for wines} &= \frac{(13)^2 + (11)^2 + \cdots + (7)^2}{4} - \text{c.t.}\\ &= 200 - 180\\ &= 20\\ \text{Total s.s.} &= 4(1^2 + 2^2 + 3^2 + 4^2 + 5^2) - \text{c.t.}\\ &= 4(55) - \text{c.t.}\\ &= 220 - 180\\ &= 40\end{aligned}$$

Then

$$W_c = \frac{20 - \frac{1}{4}}{40 + \frac{2}{4}} = \frac{19.75}{40.5} = 0.488$$

and

$$F = \frac{3(0.488)}{1 - 0.488} = \frac{1.464}{0.512} = 2.86$$

with

$$\text{d.f. for numerator} = (5 - 1) - \frac{2}{4} = 3.5$$

$$\text{d.f. for denominator} = 3(3.5) = 10.5$$

From the F-table the value of F at the 5 per cent level is estimated as 3.54. The calculated value is therefore not significant, and it may be concluded with a fair degree of confidence that the judges do not exhibit a noticeable degree of agreement in their ranking of the wines. It is not appropriate, therefore, to estimate an over-all ranking for the wines based on the combined estimates of the judges.

Setting Up the Taste Experiment

The experimental design to be used in any taste test is determined by the conditions of the experiment—that is, what answer is to be derived. Is it to be a difference test for quality control or a test of discrimination, or is preference involved?

The paired-sample test is well adapted to all preference and difference tests where the quality may be stated without ambiguity. Questions such as "Which do you prefer?", "Which is sweeter?", and "Which contains SO_2?", etc., may be answered statistically with this design.

For example, two wines, A and B, are to be compared for a detectable difference in sweetness, and a chemical test has shown A to be sweeter than B. Let us assume a panel of 20 tasters.

The first step is to line up 40 wine glasses in pairs, one pair for each taster. Next the glasses must be coded. Since certain numbers may introduce bias in the judgment, code numbers are picked at random from a book of random numbers. Two-digit numbers are sufficient for coding; such a set may be found in Fisher and Yates (1953). From any page of the two-digit numbers pick a starting point at random, then start reading numbers (vertically or horizontally) until 40 numbers have been taken. The numbers may be, for example, 55, 23, 32, 17, 49, 84, etc. These are arranged in pairs, 55–23, 32–17, 49–84, etc., and are marked on the 20 pairs of glasses with a wax pencil.

The next step is to determine which wine (A or B) is to be poured into the first glass of each pair. This may be done by writing numbers from 1 to 20 on 20 pieces of paper, placing them in a bowl, stirring thoroughly, and then drawing them out blindly, one by one, until all of the numbered papers

have been taken from the bowl. The number drawn first determines the pouring of wine into set 1, the second for set 2, etc. If the number drawn for any particular set is even, wine A is to be poured into the first glass of that set, with wine B into the second glass. If the number is odd, the reverse order is poured. In this way, 10 of the sets will contain wine A in the first glass and the other 10 will contain wine B in the first glass. This method ensures randomness in the wine serving. When the wine has been poured, it may be presented to the 20 members of the panel, one set to a member, with an appropriate score card. An example of such a card is given in table 13. Instructions are usually given to taste the samples from left to right.

TABLE 13

EXAMPLE OF SCORE CARD FOR PAIRED SAMPLE TEST

Judge.. Date..

Test No. ....

Please check below which is the sweeter sample.

Sample	Sweeter
55	
....	....
23	
....	....

You must make a choice, even if only a guess.

After the tasting, the cards are tallied for the number of correct responses. Suppose that 14 are correct and 6 incorrect; we may then use the χ^2 distribution based upon a chance probability of ½ or table 2. $\chi^2 = (\,|\,X_1 - X_2\,| - 1)^2 / N$ where X_1 = number of correct = 14; X_2 = number incorrect = 6; N = total judgments = 20; then $\chi^2 = (|\ 14 - 6\ | -1)^2\ /\ 20 = (7)^2\ /\ 20 = 2.45$. Since this is a one-tailed test, a χ^2 = 2.71 is necessary to show significance at the 5 per cent level. However, the value obtained in this test, 2.45, does not quite equal the 5 per cent value and, therefore, it is to be concluded that there is not sufficient evidence to demonstrate a difference in sweetness between the two wines.

In the case described above only one set of samples was given to each taster. For economy's sake, as many sets could be given to each taster as would not cause fatigue or decreasing taste interest, but, as indicated on p. 519, seldom more than five pairs.

If the paired-sample test is to be used for the determination of preference, the coding of glasses and assignment of wine is the same as described for difference testing, with a suitable score card related to a preference. It must be remembered, however, that the test for preference is dependent upon a two-sided argument, or a two-tailed test, so that the χ^2 values for the test of significance are $\chi^2 = 3.84$ for the 5 per cent level and $\chi^2 = 6.64$ for the 1 per cent level, or use table 2.

The duo-trio test is especially adaptable to cases of quality control, where a standard of reference is required. It may also be used for difference testing where an estimate of quality is not involved. This test uses a set of three glasses, one of which is marked "S" for the reference standard or control. The other two glasses constitute a pair, one of which will contain the same wine as glass "S" and the other glass containing the other wine. First the glasses are lined up in sets of three and then a randomized coding system is applied. The first glass of each set will be marked "S," and the other two glasses, constituting a pair, will be coded as described for the paired-sample test.

TABLE 14

EXAMPLE OF SCORE CARD FOR DUO-TRIO TEST

Judge.. Date..

Test No. ....

You have three samples.

The glass marked "S" is the reference standard. Please check below which of the coded pair is *DIFFERENT* from the standard "S."

Sample	Different
55	
....	....
23	
....	....

You must make a choice, even if only a guess.

The next step is to determine which wine goes into which glass. Let us take a specific example. We have a taste panel of 20 members so that 20 sets of glasses will be necessary. We have wine A and wine B to be judged on a difference basis. As before, 20 pieces of paper, marked from 1 to 20, are placed in a bowl and mixed thoroughly. As the numbers are drawn, each glass "S" is identified with one of the two wines; in other words, if an odd number is drawn, this particular glass "S" is to contain wine A, and if the number is even, it is to contain wine B.

After this drawing, the papers are returned to the bowl and drawn again to determine which wine is to go into each glass of a pair. This procedure is identical to that used for the paired-sample test.

When the wine has been poured, each set consists of a standard glass, "S" containing wine A or wine B, and a numbered pair, one containing wine A and the other wine B, the order having been determined by randomization. Each of the judges is presented one of the sets with an appropriate score card. Such a card is shown by table 14.

After the tasting is finished and all of the score cards are in, the results are tallied for correct judgments. The χ^2 analysis is the same as that for the paired-sample test, since the probability is ½ for giving a correct response by chance. Since this test is one-tailed, being a difference discrimina-

tion, the χ^2 values to test the significance of the results will be 2.71 for the 5 per cent level and 5.41 for the 1 per cent level (see also table 3). The judges may be allowed to taste more than one set at one sitting, provided no loss of efficiency is noted.

The triangular test is comparable to the duo-trio in its usefulness in difference tasting when the quality has not been specified. It has the advantage of being somewhat stronger statistically, since the chance probability is ⅓.

Fundamentally, the triangular design consists of a set of three samples, two of which contain the identical wine, the other (odd) sample containing

TABLE 15

EXAMPLE OF SCORE CARD FOR TRIANGULAR TEST

Judge.. Date..

Test No. ....

Two of these samples are identical and one is odd. Please check the *odd* sample.

Sample	Odd
55	....
23	....
32	....

You must make a choice, even if only a guess.

a different wine. The taster is asked to designate the odd sample. If the two wines are wine A and wine B, the six possible arrangements of the three glasses are AAB, ABA, BAA, BBA, BAB, and ABB. In setting up the experiment, a random distribution is made of these possible presentations so as to control any positional bias. This is done by placing six pieces of paper, numbered from 1 to 6, in a bowl. The numbers may be assigned to the arrangements: AAB = 1, ABA = 2, BAA = 3, BBA = 4, BAB = 5, and ABB = 6. The numbers are drawn one at a time, and the corresponding arrangement is given to the sample sets sequentially. When the six numbers are exhausted, if more numbers are needed the operation is repeated. The three glasses should be coded by the procedure described for the paired-sample test.

When the coding and pouring of the samples has been completed, these are presented to the tasters, one to a judge. A typical score card is shown in table 15.

On completion of the tasting, the cards are tabulated for correct judgments. The test for significance is $\chi^2 = (|4X_1 - 2X_2| - 3)^2 / 8N$ where X_1 = correct judgments, X_2 = incorrect judgments, and N = total number of judgments. See also table 4.

Consumer-Preference Studies

Few reliable consumer-preference studies have been conducted with California wines, and these have nearly all been made at the California State Fair at Sacramento and the Fresno District Fair. The results may or may not apply to all parts of the adult California population. It is less likely that they would apply to groups in other states. The fairs were used because they offer sufficient attendance over a period of time to permit taste testing. As an example of the problems encountered in one year, 17,538 responses were obtained, of which 5 per cent were rejected for incomplete information.

The artificiality of the conditions prevalent at a fair is also a factor. In addition, it is very difficult to determine whether or not the present

TABLE 16

EXAMPLE OF HEDONIC SCALING FOR PAIRED COMPARISONS

SAMPLE 1	SAMPLE 2
Like Extremely...........	Like Extremely...........
Like Very Much...........	Like Very Much...........
Like Moderately...........	Like Moderately...........
Like Slightly...........	Like Slightly...........
Neither Like nor Dislike...........	Neither Like nor Dislike...........
Dislike Slightly...........	Dislike Slightly...........
Dislike Moderately...........	Dislike Moderately...........
Dislike Very Much...........	Dislike Very Much...........
Dislike Extremely...........	Dislike Extremely...........

preference of the tasters will be reflected in their preference after further exposure to wines. There is practically no information on the learning sequence of wine drinkers, if any. For some of the problems in this field see Filipello (1956*b*). One thing is clear: small-panel test results cannot be substituted for consumer acceptance, as Harries (1953) indicated. For some of the advantages of consumer testing, see Nair (1949). For methods of selecting consumer panels see Anon. (1958) and Filipello, *et al.* (1958). The latter reported a much better response to their mailed questionnaires as the economic level became higher; the type of letter and the recruiting agency were also important. With their panels general usage of wine increased as the age of the consumers advanced and as the number of adults in a family became larger. Parental and extrafamilial influences also affected wine consumption.

Preference tests are usually related to mass-panel treatments, consideration of which is beyond the scope of this publication. However, some forms of hedonic scaling have been well received for testing preferences. One form of this is the dual hedonic scale for paired-sample comparisons, given by table 16. With this form of scaling, two wines may be compared directly and the preference assessed by statistical means. The usual precautions are taken in the serving order of the samples so as to eliminate positional bias.

One method, however, would be suitable for small panels of expert judges, and this method is applicable when *two or more* wines are to be judged for

preference. This is a method of preference by paired comparison. The details of this method and its analysis are given by Scheffé (1952). Essentially, all samples of wine are compared with each other in all combinations, two at a time, using both orders of serving. In this way, a preference parameter may be derived to represent each wine. The score card employed is given in table 17.

Sensory and Chemical Analyses

It would be very desirable to replace organoleptic examination with chemical analyses. Some progress has already been made. Wines with excessive volatile acidity or sulfur dioxide are always of inferior quality. Definite relations exist between a number of chemical characteristics of the wine and its quality.

TABLE 17

EXAMPLE OF SCORE CARD FOR PAIRED COMPARISONS

I prefer sample 1 over sample 2 EXTREMELY	....
I prefer sample 1 over sample 2 VERY MUCH	....
I prefer sample 1 over sample 2 MODERATELY	....
I prefer sample 1 over sample 2 SLIGHTLY	....
I have no preference	....
I prefer sample 2 over sample 1 SLIGHTLY	....
I prefer sample 2 over sample 1 MODERATELY	....
I prefer sample 2 over sample 1 VERY MUCH	....
I prefer sample 2 over sample 1 EXTREMELY	....

Baker and Amerine (1953) showed that tasting quality could be roughly predicted by using selected analytical data in predicting equations. For white table wines they employed total acidity, volatile acidity, sulfur dioxide, aldehyde, pH, and color in the predicting equation. The squares of the multiple correlation coefficients between the scores predicted by estimating equations and the scores given by the five tasters were 0.546, 0.737, 0.761, 0.336, and 0.554. For red table wines they employed alcohol, total acid, pH, tannin, and volatile acidity in the equation. The squares of the multiple correlation coefficients were 0.643, 0.322, 0.280, 0.362, and 0.641. The poorer values for the red wine were explained by the fact that a Cabernet wine was used, and the varietal aroma is an important factor in its quality. No analytical procedure for determining the amount of Cabernet aroma is available, so this important factor could not be included in the predicting equation. Baker (1954) later studied these data by factorial analyses. His findings were in close agreement with those resulting from the analysis of variance procedure used in the earlier study. In a factorial analysis there emerges for each type of wine a general quality factor, not apparent intuitively, which depends primarily on certain, perhaps different, variables in each case and which is affected only slightly by other variables.

It may be noted here that should complete analytical data be obtainable for a wine and the accuracy of the predicting equation not be diminished by omitting the data for a component, this would be presumptive evidence that the component was not of critical importance to the organoleptic quality rating. So far as we know, this has not been attempted.

There are many difficulties in replacing organoleptic examination with chemical analyses, however. In the first place, our knowledge of the influence the chemical constituents have on odor is still very incomplete. Except for non-*Vitis vinifera* varieties, such as Concord, we have little information on the compound (or compounds) responsible for the unique aroma of varieties. This is particularly important with varieties such as Cabernet Sauvignon, Gewürztraminer, Sauvignon blanc, or the several muscats.

Second, organoleptic examination is a balancing of many single impressions. The relative pleasantness of the various constituents of the wine is desired. In making decisions such as this the mind seems to act, knowingly or unknowingly, like a vast calculator. The predominance, absence, or sufficiency of many compounds is synthesized into a quality evaluation. Perhaps one day we shall be able to do this by complicated predicting equations. This might be feasible for simple wines that lack a varietal or aged odor. Such wines are apt to have a narrow quality range anyway, and relatively simple predicting equations could be developed for evaluating their quality from analytical data. But for wines of high quality, especially when aged until they develop a bouquet, predicting equations still seem somewhat distant. Ribéreau-Gayon (1950) speaks of the impossibility of replacing tasting by chemical analyses because it is impossible to relate analytical data to organoleptic impressions. Admittedly this is difficult and at present, for certain odors, impossible. But we should not overlook the possibility of doing so in the future.

The sensitivity of the olfactory sense has already been mentioned. Thus, the taster often distinguishes odorous components which can be measured only by rather delicate chemical analyses. Examples might be hydrogen sulfide, hydrogen cyanide, mercaptans, and the trace compounds responsible for the fruit flavors.

Analytical data are also valuable to the taster during his training period. The inexperienced taster gains considerable confidence when he finds his organoleptic result confirmed by analytical data. Likewise, if he finds that he has not perceived a detectable difference, he will be encouraged to retaste and to find the difference.

One of the challenges of quality evaluation of natural products is its subjective nature. It is not likely that objective chemical determinations and predicting equations will soon eliminate the factor of personal preference and experience as a method of evaluating quality wines. In this publication we have tried to give objective procedures which will help to reduce uncertainty in measuring the reliability of the judges and their scores. Moreover, a knowledge of sensory tasting should add to one's enjoyment of wines as a beverage.

ACKNOWLEDGMENTS

We are grateful to the late Professor J. G. B. Castor, to Professors J. F. Guymon and A. D. Webb, and to Mr. Harold Berg for reading the manuscript and making suggestions for its improvement. The errors that remain are those of the authors. Suggestions for improvement of future editions would be gratefully received.

LITERATURE CITED

ALDER, H. L., and E. B. ROESSLER

1951. Introduction to probability and statistics. Edwards Bros., Inc., Ann Arbor, Michigan. 136 p.

1954. Statistical procedures. University of California, Davis, vii, 436 p. (Mimeographed)

ALLEN, H. WARNER

1931. The romance of wine. Ernest Benn, Limited, London. 264 p. (See p. 225–46.)

AMERINE, M. A.

1948*a*. An application of "triangular" taste testing to wines. The Wine Review **16**(5): 10–12.

1948*b*. Hydroxymethylfurfural in California wines. Food Res. **13**:264–69.

1954. Composition of wines. I. Organic constituents. Advances in Food Research 5:353–510.

1958. Unpublished data. Department of Viticulture and Enology, Davis, California.

AMERINE, M. A., and E. FEDUCHY

1954. Los resultados de la cata del vino y del análisis químico. Bol. Inst. Nac. Invest. Agron. **14**(31):353–75.

AMERINE, M. A., and M. A. JOSLYN

1951. Table wines. The technology of their production in California. University of California Press, Berkeley and Los Angeles. 397 p. (See p. 187–95.)

AMERINE, M. A., C. S. OUGH, and C. B. BAILEY

1959. Color values of California wines. Food Tech. **13**:170–75.

AMERINE, M. A., and E. B. ROESSLER

1952. Techniques and problems in the organoleptic examination of wines. Proc. Amer. Soc. Enol., 1952, p. 97–115.

AMERINE, M. A., and G. THOUKIS

1958. The glucose-fructose ratio of California grapes. Vitis **1**:224–29.

ANON.

1958. Flavor research and food acceptance. Reinhold Publishing Corporation, New York, vi, 391 p.

BAKER, G. A.

1954. Organoleptic ratings and analytical data for wines analyzed into orthogonal factors. Food Res. **19**:575–80.

BAKER, G. A., and M. A. AMERINE

1953. Organoleptic ratings of wines estimated from analytical data. Food Res. **18**:381–89.

BAKER, G. A., M. A. AMERINE, and E. B. ROESSLER

1954. Errors of the second kind in organoleptic difference testing. Food Res. **19**:206–10.

BATEN, W. D.

1947. Materials for removing taste effects in organoleptic tests. J. Home Econ. **39**:30–32.

BEATTIE, G. B.

1949. Taste test techniques as applied to beer and soft drinks. Bottling **102**:20, 22, 24, 28, 32, 34, 36, 38.

BENGTSSON, K., and E. HELM

1946. Principles of taste testing. Wallerstein Laborat. Comm. **9**:171–80.

BERG, H., and A. D. WEBB

1955. California wine types. University of California, Davis. 20 p. (Mimeographed).

BERG, H. W., F. FILIPELLO, E. HINREINER, and A. D. WEBB

1955*a*. Evaluation of thresholds and minimum difference concentrations for various constituents of wines. I. Water solutions of pure substances. Food Tech. **9**:23–6.

1955*b*. Ibid. II. Sweetness: the effect of ethyl alcohol, organic acids and tannin. Ibid: 138–40.

BLAHA, J.

1948. Personal communication.

Boggs, M. M., and H. L. Hanson
1949. Analysis of foods by sensory difference tests. Advances in Food Research 2:219–58.
Boggs, M. M., and A. C. Ward
1950. Scoring technique for sulfited foods. Food Tech. 4:282–84.
Boring, E. G.
1942. Sensation and perception in the history of experimental psychology. D. Appleton-Century Co., New York. 644 p.
Bradley, R. A., D. B. Duncan, and M. E. Terry
1950–1953. Statistical methods for sensory difference tests of food quality. Virginia Agr. Exp. Sta., Blacksburg. (Bi-annual reports to USDA, Bur. Agr. Econ. Nos. 1–6).
Brenner, M. W., J. L. Owades, and T. Fazio
1955. Determination of volatile sulfur compounds. IV. Further notes on mercaptans. Proc. Amer. Soc. Brew. Chem., 1955, p. 125–32.
Brown, E. M.
1950. A new off-odor in sweet wines. Proc. Amer. Soc. Enol., 1950, p. 110–12.
Brunet, R.
1908. Dégustation des vins. Revue Vitic. 30:384.
1941. Le goût de terroir des vins. Le Moniteur Vinic. 86(30):57.
n.d. La dégustation des vins. Bureaux du Moniteur Vinicole, Paris. 34 p.
Byer, A. J., and D. Abrams
1953. A comparison of the triangular and two-sample taste-test methods. Food Tech. 7:185–187.
Byer, A. J., and P. P. Gray
1953. Some considerations in applying systematic taste testing to beer. Wallerstein Laborat. Comm. 16:303–14.
Cairncross, S. E., and L. B. Sjostrom
1950. Flavor profiles—a new approach to flavor problems. Food Tech. 4:308–11.
Cameron, A. T.
1947. The taste sense and the relative sweetness of sugars and other sweet substances. Sugar Research Foundation, Inc., New York. 72 p.
Castor, J. G. B.
1954. Fermentation products and flavor profiles of yeasts. Wines and Vines 35(8):29–31.
Chamberlain, N. W.
1955. A general theory of economic process. Harper, New York. 370 p. (See p. 181–82.)
Chauvet, J.
1950. L'arôme des vins fins. Bull. Inst. Natl. Appel. Orig. Vins et Eaux-de-Vie 43:8–17.
Cordonnier, R.
1955. Recherches de l'addition frauduleuse d'aromatisants aux vins doux naturels; observations sur le parfum naturel de ces vins. Compt. Rend. Acad. Agr. France 41:399–403.
1956. Recherches sur l'aromatisation et le parfum des vins doux naturels et des vins de liqueur. Annales Technol. Agr. 5:75–110.
Cover, S.
1936. A new subjective method of testing tenderness in meat. Food Res. 1:287–95.
Crawford, C., R. J. Bouthilet, and A. Caputi, Jr.
1958. Color standards for white wines. Amer. Jour. Enol. 9:194–201.
Crocker, E. C., and F. L. Henderson
1927. Analysis and classification of odors. Amer. Perfumer 22:3–10.
Crowther, R. F.
1951–1952. Flavours and odours from yeasts. Report Hort. Prod. Lab. Vineland, Ontario, 1951–1952. p. 80–83.
Cruess, W. V.
1935. Suggested scoring system for wines. Fruit Prod. Jour. 14:269–270.
Duncan, D. B.
1955. Multiple range and multiple *F* tests. Biometrics 11(1):1–42.

ELSBERG, C. A.

1935. The sense of smell. VIII. Olfactory fatigue. Bull. Neur. Inst. New York **4**:479–95.

1937. *Ibid.* XIV. The relation of the cerebral cortex to the olfactory impulse and the areas in the brain involved in fatigue of the sense of smell. *Ibid.* **6**:118–25.

ENGEL, R.

1933. Sur l'art de déguster et de présenter les grands vins. Rev. Vitic. **78**:261–66.

ENGEL, R., and R. PERNOT

1937. L'acide carbonique dans le dégustation des grands vins. Rev. Vitic. **87**:315–23.

EYSENCH, H. J.

1941. A critical and experimental study of color preferences. Amer. Jour. Psychol. **54**:385–94.

FABIAN, F. W., and H. B. BLUM

1943. Relative taste potency of some basic food constituents and their competitive and compensatory action. Food Res. **8**:179–93.

FILIPELLO, F.

1956*a*. A critical comparison of the two-sample and triangular binomial designs. Food Res. **21**:235–41.

1956*b*. Factors in the analysis of mass panel wine-preference data. Food Tech. **10**:321–26.

1957. Organoleptic wine-quality evaluation. II. Performance of judges. Food Tech. **11**:51–53.

FILIPELLO, F., H. W. BERG, and A. D. WEBB

1958. A sampling method for household surveys. I. Panel recruitment for testing wines. II. Panel characteristics and their relation to usage of wine. Food Tech. **12**:387–90; 508–10.

FISHER, R. A., and F. YATES

1953. Statistical tables for biological, agricultural, and medical research. 4th ed. Hafner Publishing Co., Inc., New York. 126 p.

GALLAY, R., and L. BENVEGNIN

1950. Les enseignements d'une dégustation. Rev. Romande d'Agric., de Vitic. et d'Arbor. **6**:38–39.

GELDHARD, F. A.

1953. The human senses. John Wiley & Sons, New York. 365 p.

GENTILINI, L.

1949. L'anidride solforosa nei vini. Riv. Viticolt. ed Enol. (Conegliano) **2**:243–244.

GÖPFERT, W.

1953. Über den Nachweis von weinfremden Aromastoffen im Dessertwein. Deut. Lebensm.-Rund. **49**:86–8.

GOETZL, F. R.

1949. Studies on human appetite. Proc. Wine Tech. Conf. Davis, Calif., 1949, p. 34–47. (See also Jour. Applied Physiol. **2**:619–26. 1950; and **4**:30–36. 1951.)

GOT, N.

1953. La dégustation des vins. Sodiep, Béziers. 157 p.

GRAZZI-SONCINI, G.

1892. Wine. Classification; wine tasting; qualities and defects. Translated by F. T. Bioletti. State Printing Office, Sacramento. 56 p. (Appendix E to the Biennial Report of the Board of State Viticultural Commissioners for 1891–92).

GRIDGEMAN, N. T.

1955. Taste comparisons—two or three? Food Tech. **9**:148–50.

GROHMANN, H., and F. H. MUHLBERGER

1954. Über die Unterscheidung zwischen natürlichen und zugesetzten Mengen an Vanillin in Dessertweinen und Aperitifs. Deut. Lebensm.-Rund. **50**:183–86.

GUILFORD, J. P.

1954. Psychometric methods. McGraw-Hill Book Co., Inc., New York. ix, 597 p.

GUYMON, J. F., and J. E. HEITZ

1952. The fusel oil content of California wines. Food Tech. **6**:359–62.

HAAGEN-SMIT, A. J., F. N. HIROSAWA, and T. H. WANG

1949. Chemical studies on grapes and wines. I. Volatile constituents of Zinfandel grapes (*Vitis vinifera*, var. Zinfandel). Food Res. **14**:472–80.

HALLGARTEN, S. F.
1951. Rhineland wineland. Paul Elek, London. 199 p. (See p. 162.)

HARRIES, J. M.
1953. Sensory tests and consumer acceptance. Jour. Sci. Food Agr. **4**:477–82.
1956. Positional bias in sensory assessments. Food Tech. **10**:86–90.

HARRISON, S., and L. W. ELDER
1950. Some applications of statistics to laboratory taste testing. Food Tech. **4**:434–39.

HELM, E., and B. TROLLE
1946. Selection of a taste panel. Wallerstein Lab. Comm. **9**:181–94.

HENNIG, K., and F. VILLFORTH
1942. Die Aromastoffe der Weine. I. II. Vorratspflege und Lebensmittelforschung **5**:181–99, 313–33.

HENNING, H.
1924. Der Geruch. 2. Aufl. Barth, Leipzig. 434 p.

HINREINER, E.
1956. Organoleptic evaluation by industry panels—the cutting bee. Food Tech. **10**: 203–5.

HINREINER, E., F. FILIPELLO, H. W. BERG, and A. D. WEBB
1955*a*. Evaluation of thresholds and a minimum difference concentrations for various constituents of wines. IV. Detectable differences in wines. Food Tech. **9**:489–90.

HINREINER, E., F. FILIPELLO, A. D. WEBB, and H. W. BERG
1955*b*. *Ibid.* III. Ethyl alcohol, glycerol and acidity in aqueous solution. *Ibid.* **9**:351–53.

HOLLEY, R. W., B. STOYLA, and A. D. HOLLEY
1955. The identification of some volatile constituents of Concord grape juice. Food Res. **20**:326–31.

HOPKINS, J. W.
1946. Precision of assessment of palatability of foodstuffs by laboratory panels. Canad. Jour. Res. **24F**:203–14.
1953. Laboratory flavor scoring: two experiments in incomplete blocks. Biometrics **9**:1–21.

ISHLER, N. H., E. A. LANE, and A. J. JANISCH
1954. Reliability of taste testing and consumer testing methods. II. Code bias in consumer testing. Food Tech. **8**:389–91.

JONES, L. V., D. R. PERYAM, and L. L. THURSTONE
1955. Development of a scale for measuring soldiers' food preferences. Food Res. **20**:515–20.

JUNGENDFEIN, K.
1937. Fachausdrücke für Geruchs- und Geschmacksempfindungen bei der Sinnenprobe von Weinen. Das Weinland **9**(1):4–6.

KENDALL, M. G.
1948. Rank correlation methods. Griffin, London. 160 p.

KEPNER, R. E., and A. D. WEBB
1956. Volatile aroma constituents of *Vitis rotundifolia* grapes. Amer. Jour. Enol. **7**:8–18.

KIELHÖFER, E.
1949. Die Beurteilung der organoleptischen Eigenschaften des Weines und ihre Auswertung, insbesondere bei kellertechnischen Versuchen. Der Weinbau. Wissenschaftliche Beihefte **3**:262–71.

KLENK, E.
1950. Die Weinbeurteilung nach Farbe, Klarheit, Geruch und Geschmack des Weines. Eugen Ulmer, Stuttgart. 118 p.

KOCH, J.
1955. Über die von der Deutschen Landwirtschaftsgesellschaft (DLG) veranstaltete dritte DLG. Leistungsprüfung für Süssmoste. Schweiz. Zeit. f. Obst- und Weinbau **64**:219–24. (See also Die Qualitätsbeurteilung der Fruchtweine. Die Obst- und Gemüseverwertungsindustrie **39**:330–31, 1954.)

KRUM, J. K.
1955. Sensory panel testing. Food Engin. **27**(7):74–83.

KUTTER, F.
1948. Beurteilung des Biergeschmackes in Ziffern. Schweiz. Brau-Rundschau **59**:108–9.

Lane, E. A., N. H. Ishler, and G. A. Bullman
1954. Reliability of taste testing and consumer testing methods. I. Fatigue in taste testing. Food Tech. 8:387–88.
Lockhart, E. E.
1951. Binomial systems and organoleptic analysis. Food Tech. 5:428–31.
MacKinney, G., and C. O. Chichester
1954. The color problem of foods. Advances in Food Research 5:301–51.
Marescalchi, A.
1949. La degustazione e l'apprezzamento dei vini. Casa Editrice S.A. Fratelli Marescalchi, Casale Monferrato. viii, 153 p.
Marteau, G.
1953. Recherche de la qualite par l'examen organoleptique des vins. Prog. Agr. et Vitic. 140:281–89, 310–13.
Mathieu, L.
1902. Le bouquet des vins. Rev. Vitic. 18:263–67.
1911. Degustation et analyse chimique des vins. Rev. Vitic. 35:185–89, 213–18, 284–86.
Mensio, C.
1957. Manuale sull'assaggio e l'apprezzamento dei vini. Ordine Nazionale degli Assaggiatori di Vino, Asti. 55 p.
Moncrieff, R. W.
1951. The chemical senses. John Wiley and Sons, Inc., New York. 468 p.
Morrow, A. R.
1936. Wine tasting as a profession. Wines and Vines 17(5):14.
Mullins, L. J.
1955. Olfaction. Annals New York Acad. Sci. 62:247–76.
Nair, J. H.
1949. Mass taste panels. Food Tech. 3:313–16.
Office International du Vin
1959. Lexique vitivinicole international. In press.
Optical Society of America, Committee on Colorimetry
1953. The science of color. Thomas Y. Crowell Co., New York. viii, 385 p.
Palieri, G., and P. G. Garoglio
1947. Il valore commerciale del vino e l'analisis organoletica. Alim., Prod. Agr. e Sost. Agr. 3:9–10.
Papakyriakopoulos, V. G., and M. A. Amerine
1956. Sensory tests on two wine types. Amer. Jour. Enol. 7:98–104.
Parker, G. H.
1922. Smell, taste and allied senses in the vertebrates. J. P. Lippincott Co., Philadelphia. 192 p.
Paul, T.
1917. Physikalische Chemie der Lebensmittel. IV. Wissenschaftliche Weinprobe zur Festsellung des Beziehungen zwischen der Stärke des sauren Geschmackes und der Wasserstoffionenkonzentration. Zeit. f. Elektrochemie 23:87–93.
Peglion, V.
1900. Ueber die wahrsheinliche Ursache des Erdgeschmackes des Weines. Staz. Sperim. Agrar. Ital. 33:525–30.
Peryam, D. R.
1950. Quality control in the production of blended whiskey. Ind. Quality Control 7(3):17–21.
Peryam, D. R., and V. W. Swartz
1950. Measurement of sensory differences. Food Tech. 4:390–95.
Peynaud, E.
1937. Études sur les phénomènes d'esterification. Rev. Vitic. 86:209–15, 227–31, 248–53, 299–301, 394–96, 420–23, 440–44, 472–75; 87:49–52, 113–16, 185–88, 242–49, 278–95, 297–301, 344–50, 362–64, 383–85.
Pfaffman, C.
1954. Variables affecting difference tests. *In* Food acceptance testing methodology. Advisory Board on Quartermaster Research and Development, Committee on Foods, National Academy of Sciences, National Research Council. 115 p. (See p. 4–17.)
1956. Taste and smell. Ann. Rev. Psychol. 7:391–408.

Power, F. G.
1921. The detection of methyl anthranilate in fruit juice. Jour. Amer. Chem. Soc. 43: 377–81.
Power, F. B., and V. K. Chesnut
1921. The occurrence of methyl anthranilate in grape juice. Jour. Amer. Chem. Soc. 43:1741–42.
1923. Examination of authentic grape juices for methyl anthranilate. Jour. Agr. Res. 23:47–53.
Prostoserdov, N. N.
1948. Personal communication to J. Blaha. (See also Vinodelie i Vinogradarstvo SSSR 6(10–11):50–51. 1945.)
Puls, E.
1939. Die Weinkostprobe. Prüfung und Beurteilung von Wein. D. Meininger, Neustadt an der Weinstrasse. 24 p.
Rao, C. R.
1950. Sequential tests of null hypotheses. Sankhya **10**:361–70.
Ribéreau-Gayon, J.
1950. La dégustation. Conférence faite à Bruxelles le 26 février 1950 ... Causse, Graille, Castelnau, Montpellier. 12 p.
Robinson, W. B., N. J. Shaulis, and C. S. Pederson
1949. Ripening studies of grapes grown in 1948 for juice manufacture. Fruit Prod. Jour. 29:36–7, 54, 62.
Roessler, E. B., J. Warren, and J. F. Guymon
1948. Significance in triangular taste tests. Food Res. **13**:503–05.
Rouleau, H.
1953. La dégustation de la bière et les dégustateurs. Le Petit Jour. du Brasseur **61**: 177–80, 194–97.
Sale, J. W., and J. B. Wilson
1926. Distribution of volatile flavor in grapes and grape juices. Jour. Agr. Res. 33: 301–10.
Satorius, M. J., and A. M. Child
1938. Problems in meat research I. II. Food Res. 3:627–35.
Schanderl, H.
1950. Die Mikrobiologie des Weines. Eugen Ulmer, Stuttgart. 211 p.
Scheffé, H.
1952. An analysis of variance for paired comparisons. J. Amer. Stat. Assn. **47**:381–400.
Scott, R. D.
1923. Methyl anthranilate in grape beverages in flavors. Ind. and Eng. Chem. **15**: 732–33.
Skramlik, E. von
1926. Handbuch der Physiologie der niederen Sinne. Georg Thieme, Leipzig. 532 p.
Slater, R. R., and G. K. Frenkel
1948. A study of the influences of certain variables in taste testing applied to beer. Proc. Amer. Soc. Brew. Chem., 1948, p. 101–12.
Snedecor, G. M.
1946. Statistical methods—applied to experiments in agriculture and biology. 4 ed. Iowa State College Press, Ames, Iowa. 485 p.
Stone, I.
1954. Color methods in the brewing industry. ASTM Bull. **201**:40–42.
Thurstone, L. L.
1952. The measurement of values. The Psychometric Laboratory, Univ. of Chicago, No. **77**. (See also No. **68**, 1951, and No. **52**, 1950.)
Tilgner, D. J.
1957. Analiza organoleptyczna zywnosci. Wyd. Przemyslu Lekkiego i Spozywczego, Warszawa. 364 p.
Troost, G., and E. Wanner
1955. Weinprobe. Weinansprache. 3 Aufl. Verlag Sigurd Horn, Frankfurt/Main. 63 p.

VENTRE, J.
1931. Traité de vinification pratique et rationnelle. II. Le vin. Librairie Coulet, Montpellier. 487 p. (See p. 107–15.)
1935. Les levures en vinification. Coulet et Fils, Montpellier, 59 p. (See also Prog. Agr. et Vitic. **106**:111–15, 135–40, 153–55, 183–87, 1936.)

VIRDEN, C. J.
1949. Interpretation of tasting experiments. Jour. Inst. Brew. **55**:228–33.

WAGNER, K. G.
1950. Sprachliche Grundlage der Sinnenprüfung von Lebensmitteln. VI. Geruch und Geschmack der Lebensmittel. Zeit. f. Lebensmittel-Untersuch.- u. Forschung **90**: 36–46.

WALD, A.
1947. Sequential analysis. John Wiley and Sons, Inc. New York. 212 p.

WEBB, A. D., and R. E. KEPNER
1957. Some volatile aroma constituents of *Vitis vinifera* var. Muscat of Alexandria. Food Res. **22**:384–95.

WINE INSTITUTE
1956. California wine type specifications. Revised May 2, 1956. San Francisco, Calif. 8 p. (mimeo.)

WINKLER, A. J., and M. A. AMERINE
1937. Color in California wines. I. Methods for measurement of color. Food Res. **3**: 429–38.

WOOD, E. C.
1949. Organoleptic tests in the food industry. III. Some statistical considerations in organoleptic tests. Jour. Soc. Chem. Ind. **68**: 128–31.

YULE, G. U., and M. G. KENDALL
1950. An introduction to the theory of statistics. 14 ed. Harper Publishing Co., New York. 701 p.

ZASLAVSKIĬ, A. S.
1955. The mousy flavor in wine (transl.). Sadovodstvo vinogradarstvo i vinodelie Moldaviĭ **2**:43–44.

ZELLNER, H.
1927. Weinbrand, seine Verfälschungen und deren Nachweis. Zeit. Untersuch. Lebensm. **54**:389–90.

5m-6,'59(7005)BEG